

Screening for osteoporosis

An evidence map and umbrella review for the UK National Screening Committee

Version: 4

Author: Warwick-Birmingham Screening Evidence Synthesis Group (WeBS-ESG)

Date: July 2026

This report was commissioned by the National Institute for Health and Care Research (NIHR) Evidence Synthesis Programme. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

The UK National Screening Committee secretariat is hosted by the Department of Health and Social Care

Contents

Contents.....	2
About the UK National Screening Committee	4
List of abbreviations.....	5
List of tables	6
List of figures	7
Plain English summary	8
Executive summary	9
Purpose of the review	9
Background.....	9
Focus of the review.....	9
Recommendation under review.....	10
Findings	10
Recommendations on screening.....	12
Limitations	13
Evidence uncertainties and gaps to be addressed	14
Introduction and approach	15
Background.....	15
Current policy context and previous reviews	18
Objectives	19
Evidence map.....	20
Aims of the evidence map.....	20
Search methods and results	20
Summary of findings	23
Question 1a: What is the volume and type of evidence exploring the clinical benefit of whole population screening in women over 65, in reducing osteoporotic fractures compared to standard care?	23
Question 1b: What is the volume and type of evidence exploring the clinical benefit of targeted screening of women aged under 65 with risk factors for osteoporosis, in reducing osteoporotic fractures compared to standard care?	23
Question 2: What is the volume of evidence exploring the cost-effectiveness of screening (whole population and targeted) for osteoporosis?.....	23
Umbrella review.....	24
Aims of the umbrella review	24
Decision question	24
Methods.....	24
Eligibility for inclusion in the umbrella review	24
Data extraction	24

Methods of analysis/synthesis	24
Appraisal for quality/risk of bias tool	25
Results.....	26
Included reviews.....	26
Overview of reviews	26
Assessment of reviews.....	31
Review summary.....	53
Independent assessment of the evidence base	53
Future considerations	55
Pragmatic checklist for assessing review comparability when conclusions differ.....	56
Appendix 1 — Search strategy for the evidence map	57
Appendix 2 — Abstract reporting	68
Appendix 3 — Data extraction and risk of bias assessment.....	69
Full data extraction sheet.....	69
Summary of risk of bias	69
Trials	69
Systematic reviews.....	70
Appendix 4 — Proposed checklist	71
Step 0. Identify an anchor review	72
Step 1. Establish whether reviews are comparable	72
Step 2. Conflict analysis (reviews differ despite using the same evidence).....	72
Summary output.....	73
References.....	74

About the UK National Screening Committee

The UK National Screening Committee (UK NSC) advises ministers and the NHS in the 4 UK countries about all aspects of [population screening](#) and targeted screening and supports implementation of screening programmes.

Conditions are reviewed against [evidence review criteria](#) according to the UK NSC's [evidence review process](#).

Read a [complete list of UK NSC recommendations](#).

UK National Screening Committee, Southside, 39 Victoria Street, London, SW1H 0EU

www.gov.uk/uknsc

Blog: <https://nationalscreening.blog.gov.uk/>

For queries relating to this document, please contact: uknsc@dhsc.gov.uk.

© Crown copyright 2026

You may re-use this information (excluding logos) free of charge in any format or medium, under the terms of the Open Government Licence v3.0. To view this licence, visit OGL or email psi@nationalarchives.gsi.gov.uk. Where we have identified any third-party copyright information you will need to obtain permission from the copyright holders concerned.

Published Month 20XX

List of abbreviations

AHRQ	Agency for Healthcare Research and Quality
AUC	Area under the curve
BMD	Bone mineral density
BMI	Body mass index
CI	Confidence interval
DALY	Disability-adjusted life years
DXA	Dual-Energy X-ray Absorptiometry
FE	Fixed effect
FRAX	Fracture Risk Assessment Tool
GP	General practitioner
GRADE	Grading of Recommendations Assessment, Development and Evaluation
HEN	Health Evidence Network
HR	Hazard ratio
HRT	Hormone replacement therapy
ICD	International Classification of Diseases
IQR	Interquartile range
ITT	Intention-to-treat
MOF	Major osteoporotic fracture
NICE	National Institute for Health and Care Excellence
NOGG	National Osteoporosis Guideline Group
NR	Not reported
OR	Odds ratio
PICOS	Population, Intervention, Comparator, Outcome and study design
PP	Per-protocol
PPV	Positive predictive value
RCT	Randomised controlled trial
RE	Random effect
REA	Rapid Evidence Assessment
RoB2	Cochrane Risk-of-Bias tool for randomized trials 2
ROBIS	Risk of bias in systematic reviews
RR	Relative risk
SCORE	Simple Calculated Osteoporosis Risk Estimation
SD	Standard deviation
SE	Standard error
SERMs	Selective oestrogen receptor modulators
SOF	Study of Osteoporotic Fracture
SR	Systematic review
UK NSC	UK National Screening Committee
USPSTF	US Preventive Services Task Force
VFA	Vertebral fracture assessment

List of tables

Table 1: Summary of reviews	28
Table 2: Summary of included primary studies in the reviews.....	34
Table 3: Summary of primary studies included in the reviews.....	36
Table 4: Summary of meta-analyses and trial level estimates considered in the reviews. Estimates are HR or RR (95% CI) for screening intervention versus control unless stated otherwise	44
Table 5: Summary of meta-analysis relative treatment effect results and absolute effects. Estimates are for screening intervention versus no intervention.	47
Table 6: Validation of review author reported conclusions with numerical data and meta-analysis results presented	49

List of figures

Figure 1. PRISMA flow diagram of included and excluded records.....22

Figure 1.1. Summary of risk assessment using the Cochrane Risk of Bias 2 tool (RCTs).....69

Figure 1.2. Summary of risk assessment using the ROBINS-I tool (nonrandomised studies)69

Figure 1.3. Summary of risk assessment using ROBIS70

Plain English summary

Osteoporosis is a disease that makes bones weak and more likely to break. It affects millions of people around the world and mainly affects older women. Broken bones, such as hip fractures, can cause serious health problems or even lead to death. Finding osteoporosis early through screening could help prevent many fractures and reduce the impact on people's lives and the healthcare system.

The UK National Screening Committee (UK NSC) looked at the evidence on screening women for osteoporosis in 2019 and 2025. Some studies showed that screening women over the age of 65 slightly lowered the risk of fractures. Other studies showed no benefit. Because the results did not agree, the UK NSC decided not to recommend a national screening programme for women.

The UK NSC wanted to 1) see if new research provides stronger evidence and 2) understand the reasons for these mixed results.

This updated review looked at new evidence from 2024 onwards on:

- Whether screening women over the age of 65, or younger women with risk factors, helps lower the risk of fractures compared to usual care.
- Whether osteoporosis screening is good value for money.

We found one new review study. It reported that screening in women aged 65 years or older appears to slightly reduce hip and major fractures compared with usual care. We did not find any new evidence on screening younger women with risk factors or on whether osteoporosis screening is good value for money.

This new study was included in our second review to understand mixed results reported for osteoporosis screening. We compared all relevant reviews to understand why their conclusions differed. Most were based on the same 3 studies and reported similar results. Differences in conclusions were mainly because of how important small improvements were viewed by study authors rather than a real disagreement in the evidence.

There is some evidence that screening women aged 65 years and over might reduce hip and major osteoporotic fractures, but the results should be viewed with caution because the studies were judged to have some concerns for risk of bias. However, screening was not associated with reductions in all fractures or deaths. These results should be viewed with caution. They mainly focused on women who were already more likely to break a bone, and some women receiving usual care were tested and treated outside the study setting. This makes it harder to judge how big the benefit of screening would be at a population level.

New, high quality research results are needed before policy recommendations should be revisited.

It is important to note that the UK NSC recommends screening programmes across the four UK countries. NICE guidance supports individual patient care, which uses different evidence from that needed for a national screening programme.

Executive summary

Purpose of the review

The purpose of this review is to inform the UK National Screening Committee's (UK NSC) consideration of an osteoporosis screening programme. The review consisted of 2 components:

1. **Evidence map:** to identify any new primary studies or reviews reporting on the efficacy or cost-effectiveness of population and targeted screening for osteoporosis in women published since the 2025 UK NSC evidence map.
2. **Umbrella review:** to examine the included primary studies and reviews identified in both the previous and updated UK NSC evidence map in detail and investigate what factors (methodological or otherwise) could account for differences in conclusions between the reviews.

Background

Osteoporosis is a common, progressive bone disorder that substantially increases the risk of fragility fractures, particularly in older postmenopausal women. Prevalence and fracture risk rise sharply with age, and fractures (especially hip fractures) are associated with significant disability, excess mortality, and loss of independence. The disease imposes a considerable public health and economic burden worldwide, driven by high rates of hospitalisation, long-term care needs, and ongoing healthcare use. Many osteoporotic fractures are potentially preventable through earlier identification of osteoporosis and effective treatment, making osteoporosis an important target for preventive health strategies.

Screening aims to identify individuals at high fracture risk due to osteoporosis who may benefit from treatment, typically using clinical risk assessment tools followed by bone mineral density testing where indicated. In the UK, tools such as Fracture Risk Assessment Tool (FRAX) and QFracture are widely used in clinical practice and supported by effective treatments and national guidelines. However, the UK NSC has repeatedly concluded that the evidence is insufficient to support a national population or targeted screening programme. Reviews in 2013 and 2019 found no clear reduction in fractures from systematic screening compared with usual care and identified important uncertainties around clinical benefit and cost-effectiveness. A 2025 evidence map identified additional studies and reported mixed and inconsistent conclusions across reviews, despite reliance on similar underlying trials. These ongoing uncertainties have prompted the commissioning of the current review to clarify the evidence and inform future policy recommendations.

Focus of the review

The aim of the current review is to identify evidence on the efficacy and cost-effectiveness of population and targeted screening for osteoporosis in women and to understand what factors (methodological or otherwise) could account for differences in conclusions between the reviews identified.

The **evidence map** aimed to address the following questions:

- Question 1a: What is the volume and type of evidence exploring the clinical benefit of whole population screening in women over 65, in reducing osteoporotic fractures compared to standard of care?
- Question 1b: What is the volume and type of evidence exploring the clinical benefit of targeted screening in women aged under 65 with risk factors for osteoporosis, in reducing osteoporotic fractures compared to standard of care?
- Question 2: What is the volume of evidence exploring the cost-effectiveness of screening (whole population and targeted) for osteoporosis?

A comprehensive literature search was conducted to address all 3 questions simultaneously, including searches of electronic databases, clinical trial registries, as well as examination of reference lists of included studies and relevant systematic reviews. The search focused on evidence published from 2024 onwards.

The **umbrella review** aimed to address the following question:

- What methodological factors could explain the mixed conclusions across existing reviews on the effectiveness of screening for osteoporosis in women, considering:
 - women aged over 65 (whole population)
 - women of any age with identifiable risk factors (targeted screening)?

The reviews and included primary studies identified in the 2025 evidence map and its update as part of the current review were considered in the umbrella review. Reviews and relevant primary studies were compared, focusing on the following four key domains: Scope of review, included evidence base, meta-analyses/numerical data and author conclusions/interpretation.

The methodological quality and risk of bias of included studies were assessed using established tools appropriate to study design. Systematic reviews were appraised using the risk of bias in systematic reviews (ROBIS) tool. Randomised controlled trials were assessed using the Cochrane Risk of Bias 2 tool and non-randomised studies using ROBINS-I.

Recommendation under review

The UK NSC does not currently recommend screening for osteoporosis in postmenopausal women. This position was informed by a review commissioned in 2019, which found insufficient evidence to fully address key questions relating to test accuracy, intervention effectiveness, clinical benefit from randomised controlled trials, and cost-effectiveness in a UK setting. Overall, the review concluded that important uncertainties and evidence gaps remained. In 2024, the UK NSC commissioned an evidence map to assess the volume and type of evidence on key issues in osteoporosis screening since the 2019 review found no evidence on the clinical benefit of targeted screening in women under 65 with risk factors for osteoporosis and highlighted conflicting findings across reviews of population screening in women over 65. To better understand these disparities and address remaining uncertainties, the UK NSC has commissioned the current review.

Findings

Evidence map

- Question 1a: We identified a single systematic review that aimed to review evidence on osteoporosis screening in women 65 years or older, capturing 3 SRs and 3 RCTs (SCOOP, ROSE and SOS) reporting on four fracture outcomes (all fractures, osteoporotic fractures, major osteoporotic fractures, hip fractures) and mortality.
 - The primary intention-to-treat (ITT) analyses in 2 RCTs (ROSE and SOS) showed no reduction in any fractures or mortality. SCOOP reported the only statistically significant reduction for hip fractures (hazard ratio (HR), 0.72 [95% CI, 0.59-0.89]).
 - Most pooled estimates included the per protocol-1 (PP-1) analysis of the ROSE trial. Compared with the primary ITT analysis, the PP-1 analysis produced more favourable estimates of screening effectiveness. The PP-1 analysis excluded randomised participants without sufficient information to calculate FRAX. This analysis is more susceptible to selection bias than the primary ITT analysis because it no longer reflects the original randomised comparison. Furthermore, it was conducted in a selected subgroup of participants rather than the full trial population. However, it was not possible to determine the extent to which the favourable results reflect selection bias, subgroup effects, or both.
 - Based on these pooled analyses of the RCTs, screening was associated with a reduction in two fracture outcomes - hip fractures (pooled relative risk [RR], 0.83 [95% CI, 0.73-0.93]) and major osteoporotic fractures (pooled RR, 0.94 [95% CI, 0.88-0.99]) - when compared to usual care, with absolute risk differences reported as 5 and 6 fewer fractures per 1000 participants screened, respectively. These pooled estimates should be interpreted with caution as they may be more favourable than if estimates had been based on the primary ITT analyses.
- Question 1b: We did not identify any eligible evidence for this question.
- Question 2: We did not identify any eligible evidence for this question.

Umbrella review

Six reviews on osteoporosis screening were identified (published 2003 to 2025). After excluding an early synthesis report that predates the major trials, the remaining reviews were broadly aligned in scope and drew largely on the same 3 pivotal European randomised trials (ROSE, SCOOP, SOS). All numerical outcome data from trials and meta-analyses reported across the reviews were aligned in terms of direction of effect and statistical significance. Differences in conclusions appeared to reflect how the similar meta-analysis findings of small statistically significance effects were valued, particularly the emphasis placed on relative versus absolute effects.

ROSE and SCOOP evaluated population screening strategies in women aged 65 years and over. SOS evaluated a targeted screening strategy, enrolling women with at least one clinical risk factor for osteoporosis. Despite this, the reviews pooled data across all 3 trials. Further, the interpretation of the pooled results is complicated by pragmatic study designs and low trial quality, including the uptake of dual-Energy X-ray Absorptiometry (DXA) and treatment (where indicated) in usual care arms, and differences in how participants were selected and assessed before randomisation, which may affect generalisability and dilute observed effects.

As the meta-analyses pooled all 3 pivotal trials, the combined estimates primarily reflect the effect of offering osteoporosis screening to women aged 65 years and over. SOS contributes

data from a risk-enriched subgroup, but the pooled results do not provide a distinct summary estimate for targeted screening, which must instead be interpreted from SOS alone.

At the trial level, effects were modest. SCOOP reported the only statistically significant reduction in one fracture outcome - which was hip fractures HR, 0.72 [95% CI, 0.59-0.89]. Pooling findings produced more favourable estimates as it reflects the accumulation of small effects across trials. However, pooled estimates should be interpreted cautiously. ROSE, which provides the clearest offer-of-screening comparison in an unselected population (women undergo baseline FRAX after randomisation), did not demonstrate statistically significant hip fracture reductions on an ITT basis (HR, 1.007 [95% CI 0.894, 1.136]) but favourable estimates of screening effectiveness in a per protocol PP-1 analysis (HR, 0.838 [95% CI 0.683, 1.028]). Although illustrated here for hip fractures, the PP-1 analysis consistently produced more favourable estimates than the ITT analysis across the fracture outcomes evaluated. Most meta-analyses synthesised the ROSE PP-1 analysis rather than the ITT analysis. The PP-1 analysis was restricted to women with sufficient information to calculate baseline FRAX and therefore no longer represented the effect of offering screening to all randomised participants. Restricting the analysis in this way may reduce the balance achieved through randomisation if women with complete FRAX data differ systematically from those without, producing more favourable estimates of screening effectiveness. Consequently, although the meta-analytic evidence (summarised here from the most recent review from the evidence map; Kahwati 2025) indicates screening is associated with statistically significant reductions in hip fractures (pooled RR, 0.83 [95% CI, 0.73-0.93]) and major osteoporotic fractures (pooled RR, 0.94 [95% CI, 0.88-0.99]), but not all fractures (pooled RR, 0.96 [95% CI, 0.90-1.02]) or all-cause mortality (pooled RR, 0.99 [95% CI, 0.95-1.04]), these pooled estimates should be interpreted in light of the analytical choices made in the reviews.

Evidence for targeted screening is more limited. SOS, the only trial specifically enrolling women aged 65 to 90 years with at least 1 clinical risk factor¹, did not demonstrate statistically significant reductions in all fractures (HR, 0.97 [95% CI, 0.87-1.09]), osteoporotic fractures (HR, 0.92 [95% CI, 0.82-1.03]), major osteoporotic fractures (HR, 0.92 [95% CI, 0.82-1.05]), hip fractures (HR, 0.91 [95% CI, 0.72-1.15]), or all-cause mortality (HR, 1.02 [95% CI, 0.90-1.16]). Notably, effects might be expected to be larger in a risk-enriched population; the null findings may partly reflect suboptimal participation and adherence, as well as the waiting-list design whereby control participants with indications for DXA, vertebral fracture assessment (VFA) and blood tests, were informed and advised to seek assessment through usual care, limiting separation between groups. On the other hand, effects may have been overestimated as all women in the screening arm received risk assessment as well as VFA and BMD.

Recommendations on screening

This review found no substantive disagreement in the underlying evidence base. Differences in conclusions across reviews largely reflect variation in interpretation of trial evidence rather than conflicting results.

Considering the primary evidence, there is currently insufficient evidence in quality and quantity to recommend implementation of population or targeted screening for osteoporosis. The primary ITT analyses from the 2 trials most relevant to screening unselected women aged 65 years and over (ROSE and SCOOP) showed no overall reduction in fractures or mortality, although SCOOP reported a reduction in hip fractures. Similarly, the ITT analysis from the risk-enriched

¹ a previous fracture after age 50 years, a parental hip fracture, low body weight, rheumatoid arthritis, early menopause, malabsorption syndrome, chronic liver disease, type I diabetes mellitus, or immobility

subgroup trial (SOS) showed no overall reduction in fractures or mortality. The ROSE PP-1 analysis produced more favourable results than the ITT analysis. Because the ROSE PP-1 population excluded randomised participants without sufficient information to calculate baseline FRAX, it no longer reflected the original randomised comparison and was therefore more susceptible to selection bias. Furthermore, it was conducted in a selected subgroup of participants rather than the full trial population. As most meta-analyses pooled the ROSE PP-1 analysis rather than the ITT analysis, their pooled estimates should be interpreted cautiously. The more favourable pooled estimates may therefore reflect, at least in part, the inclusion of the ROSE PP-1 analysis rather than the primary ITT analysis. Although these pooled estimates suggested small statistically significant reductions in hip fractures and major osteoporotic fractures, they showed no reductions in all fractures or all-cause mortality. The clinical significance of these findings is uncertain.

Moreover, the available evidence comes from trials with some concerns for risk of bias, and the included populations and testing lack applicability to potential screening pathways. Direct evidence supporting a targeted screening approach remains limited and biased.

Overall, the trial evidence indicates that screening may be effective for some high-risk women but the likely sub-population and fracture outcome to target cannot be identified from the current trials. Future reconsideration requires new well-designed primary trials rather than further synthesis of existing studies.

Based on the findings of this evidence map and umbrella review, no further evidence synthesis work on osteoporosis should be commissioned at the present time. The UK NSC will return to screening for osteoporosis in women in 3 years' time.

The UK NSC is the only organisation that makes recommendations on screening programmes in the four UK countries. It is important to distinguish that NICE's guideline recommendation on osteoporosis primarily support decision-making for individual patient care. The nature and strength of evidence underpinning clinical guidance differs from evidence required to justify the implementation of a nationally delivered screening programme.

Limitations

The review to identify additional evidence was conducted as an evidence map which involved methodological adjustments compared with a full systematic review. While this approach supports efficient identification of the available evidence base using a comprehensive search strategy, it provides a high-level overview of the available evidence rather than a detailed synthesis of effectiveness or study quality. The evidence map used a rapid process, including single-reviewer screening without duplicate or blinded assessment, limited full-text assessment, and no formal risk-of-bias assessment. These methods are standard for evidence maps and in this review were combined with a more formal assessment of the evidence as part of the umbrella review.

The main synthesis of evidence was based on an umbrella review of existing reviews and meta-analyses, undertaken using established systematic review methods, including predefined eligibility criteria, targeted double extraction of key outcomes, and independent quality assessment of both the included reviews and the pivotal primary trials. This assessment is therefore dependent on the scope, quality, and reporting of the available secondary evidence. The underlying evidence base remains small, comprising 3 pivotal pragmatic RCTs, which limits the certainty and generalisability of pooled estimates. In addition, heterogeneity in trial design,

recruitment strategies, and screening pathways, together with contamination of usual care arms, reduces the extent to which trial results can be directly applied to specific population or targeted screening programme models.

Evidence uncertainties and gaps to be addressed

The key uncertainty is whether the modest reductions observed in selected fracture outcomes (hip and major osteoporotic fractures) represent a clinically meaningful benefit at the population level, given the lack of consistent effects across all fractures and no evidence of reduced mortality.

Direct evidence for targeted screening in higher-risk women is limited, with only one trial specifically evaluating this approach and one trial reporting a subgroup analysis.

Additional large, well-designed randomised trials including a well-defined target population and primary screening outcome would help clarify the effectiveness of screening.

Future evidence should reflect routine clinical pathways and minimise contamination between trial arms to strengthen the evidence base.

There is a lack of evidence on the cost-effectiveness of screening (whole population and targeted) for osteoporosis.

Introduction and approach

Background

Osteoporosis is a systemic metabolic bone disorder characterized by reduced bone mass and deterioration of bone microarchitecture (1, 2). These changes increase bone fragility and, consequently, the risk of fractures (3). The most severe consequence of osteoporosis is the occurrence of osteoporotic or fragility fractures. In 2019, the global incidence of osteoporosis was estimated at 41.5 million cases, with projections suggesting 263.2 million cases between 2030 and 2034 (4). Within the European Union, approximately 5.5 million men and 22 million women are affected, resulting in about 3.5 million fractures annually (5, 6). Globally, 158 million individuals aged ≥ 50 years were at risk of osteoporotic fractures in 2010, a figure expected to double by 2040 (7). Osteoporosis is a major health concern in postmenopausal women, with fracture risk increasing sharply with age. Bone mass peaks in early adulthood before gradually declining, with loss accelerating after menopause. Epidemiological data show prevalence nearly doubling every five years, from 3.3% at ages 45–49 to 50.3% in women aged ≥ 85 years (8). Most cases in postmenopausal women are linked to oestrogen deficiency, where rapid bone loss results from increased turnover and an imbalance between resorption and formation (9).

Fragility fractures, particularly hip fractures, represent a major public health concern due to their association with significant disability and mortality (10). It is estimated that mortality from hip fractures ranges between 18 and 33% per year (5) with the mortality risk highest immediately after the fracture event and then decreasing with time (11). In 2010, osteoporosis was responsible for 5.8 million disability-adjusted life years (DALYs) lost in Europe, representing 0.83% of the global burden of non-communicable diseases (5). The DALY loss and associated mortality from osteoporosis-related fractures exceeded that of hypertensive heart disease, rheumatoid arthritis, and all cancers except lung cancer (5, 12).

The economic burden of fragility fractures is also substantial, affecting both patients and healthcare systems (13). A 2020 review reported that the cost of osteoporosis-related fracture is approximately \$17.9 and £4 billion per annum in the USA and UK respectively (14). In a 2023 study of women aged ≥ 50 years in Australia, Germany, South Korea, Spain, and the USA, those with osteoporotic fractures had significantly higher inpatient admission rates, longer hospital stays, and greater healthcare costs compared with a matched cohort of women without previous evidence of osteoporotic fractures (13). Across all countries, the adjusted cost ratios for pharmacy, inpatient, emergency, and outpatient care were consistently higher in the osteoporotic fractures cohort.

Fragility fractures can be largely prevented through the early detection, diagnosis, and treatment of osteoporosis. Implementing an effective early screening programme might help prevent fractures while reducing the clinical, humanistic, and economic burden of the disease.

Potential screening strategies

Screening for osteoporosis aims to find people who could benefit from early treatment to prevent fragility fractures and their complications. Early identification and timely referral have the potential to improve outcomes and prognosis, while good frontline diagnostic capability promotes equity and shortens waits for specialist care. Diagnosing osteoporosis early can also reduce public health costs by preventing fractures and avoiding expensive surgeries and hospital stays.

Women aged ≥ 65 years represent a key population for osteoporosis risk assessment, reflecting the strong association between increasing age and fracture risk (15). Age-based thresholds are commonly used in international guidance to identify individuals who may warrant further assessment, as a substantial proportion of first fragility fractures occur in this age group, including among women without additional clinical risk factors (16-18).

Women aged under 65 years with specific risk factors also represent an important population for risk-based assessment. Although average fracture risk is lower in younger postmenopausal women, the presence of certain clinical risk factors can substantially increase individual risk. These include low body weight or body mass index, parental history of hip fracture, cigarette smoking, excess alcohol consumption, previous fragility fracture, long-term oral glucocorticoid use, rheumatoid arthritis, and other causes of secondary osteoporosis, with some risk assessment tools also incorporating a history of falls (19). In this group, targeted assessment aims to identify women who may warrant further investigation, including bone mineral density testing, despite being below the usual age threshold for routine assessment (8). Targeted screening based on risk factors is also applicable to all age groups, where it may further refine decisions about investigation and management.

Osteoporosis assessment is based on estimating an individual's future risk of fracture and intervening when that risk exceeds a treatment threshold. In practice, this involves a stepwise approach in which clinical risk prediction tools are used to estimate 10-year fracture probability, followed by targeted use of bone mineral density (BMD) measurement to refine risk of fractures or osteoporosis and inform management decisions. BMD can be used to diagnose osteoporosis based on a T-score threshold (≤ -2.5) to prompt treatment in some settings. Alternatively, BMD scores can be incorporated into risk prediction tools to update fracture risk and inform clinical decision-making. Thus, clinical decision-making may be based on either diagnostic thresholds or estimated fracture risk.

A number of risk assessment tools have been developed to estimate the 10-year probability of a fragility fracture. The most widely used tools include the Fracture Risk Assessment Tool (FRAX) (20) and QFracture (21). Both FRAX and QFracture estimate 10-year fracture risk (absolute risk expressed as percentages) using common clinical risk factors such as smoking, alcohol use, corticosteroid exposure, and parental history of fracture. Risk estimates are used to identify individuals who require further tests such as measurement of BMD. The threshold for further testing is age dependent and increases with age. For example, a 65-year-old woman with a BMI of 25 kg/m² and no additional clinical risk factors has a FRAX 10-year major osteoporotic fracture probability of 8.4% (20), which meets the threshold for BMD testing (22). Clinical use of either tool depends on local guidelines and patient context, with caution advised in interpreting 10-year risk estimates in very elderly individuals, as these may obscure high imminent or near-term fracture risk.

In addition to these commonalities, FRAX and QFracture differ in key aspects. Unlike FRAX, QFracture includes a history of falls and a wider range of clinical variables, though it does not incorporate BMD, which FRAX allows, and which remains an important predictor of fracture risk (23). QFracture has been extensively validated in UK populations and shown to be more accurate for predicting fractures in this setting, but it is reported to be more cumbersome to use and not calibrated for other countries (21). In contrast, FRAX is widely calibrated internationally and integrated into guideline-based treatment thresholds, such as those of the National Osteoporosis Guideline Group (NOGG) in England and Wales (22).

FRAX is a computer-based algorithm (<http://www.shef.ac.uk/FRAX>) designed to estimate the 10-year probability of a major osteoporotic fracture (hip, clinical spine, humerus, or wrist) as well

as the 10-year probability of hip fracture. It calculates risk based on age, body mass index, and a set of well-validated dichotomous clinical risk factors, with the option to include femoral neck BMD for improved accuracy (24). The FRAX algorithm was developed from a series of meta-analyses incorporating data from 12 independent fracture studies conducted across North America, Europe, Asia, and Australia (25-30). Together, these studies included 60,000 men and women, with more than 250,000 person-years of follow-up, and documented over 1,100 osteoporotic hip fractures and 3,300 overall osteoporotic fractures (31). Because fracture probability varies considerably across regions, FRAX is calibrated to the local epidemiology of fracture and mortality in countries where such data are available (currently 64 countries). Since its launch in 2008, FRAX has been widely adopted worldwide, with approximately 6 million calculations performed each year in 173 countries (24).

QFracture (www.qfracture.org) is a risk assessment tool based on routinely collected electronic health record data and validated in large UK primary care populations (32). It estimates the 10-year risk of hip and major osteoporotic fractures (hip, spine, wrist) in individuals aged 30 to 85 years, without requiring BMD measurement. The algorithm incorporates a broad range of clinical risk factors, including age, sex, BMI, smoking, alcohol intake, glucocorticoid use, asthma, cardiovascular disease, history of falls, chronic liver disease, rheumatoid arthritis, type 2 diabetes, and tricyclic antidepressant use, with additional female-specific variables such as hormone replacement therapy, menopausal symptoms, parental hip fracture, gastrointestinal malabsorption, and other endocrine disorders (33). QFracture was derived from a UK prospective open cohort study of over 2 million men and women from 357 general practices (33). External validation studies report that QFracture has low positive predictive values (PPVs), with PPVs for hip fracture of approximately 4.0% at a 3% 10-year risk threshold (i.e., $\geq 3\%$ predicted risk of hip fracture over 10 years) and 4.6% at a 5% threshold in women, while negative predictive values consistently exceeded 90% (34).

Patients deemed at risk of fragility fractures after completing either FRAX or QFracture will be offered measurement of BMD in line with current UK NOGG guidelines (22). DXA is the most commonly used method to assess BMD, calculated as the amount of mineral content divided by the scanned area and reported in g/cm² and relative to a reference population (35). Site-specific BMD correlates best with fracture risk at that site but also predicts fractures elsewhere, with hip BMD considered the most accurate and clinically relevant measure (36). Measurement at both the lumbar spine and hip is recommended in all patients (37), while peripheral DXA of the radius is used when central sites are not feasible or in conditions affecting forearm BMD, such as hyperparathyroidism (38).

Osteoporosis in the UK is diagnosed using DXA-derived T-scores at the lumbar spine, hip, femoral neck, and, where indicated, the one-third (33%) radius, in line with national guidance (e.g. NOGG) (22). The T-score is the difference between the measured BMD and the BMD of a healthy young adult, reported in standard deviation units. The risk of broken bones increases by 1.5 to 2 times with each 1-point drop in the T-score with a T-score ≤ -2.5 defining osteoporosis according to World Health Organization criteria (39) and is applied in UK clinical practice (40). Modern DXA scanners can also detect vertebral fractures via lateral spine imaging (41, 42). Femoral neck BMD can be incorporated into risk prediction tools such as FRAX to refine estimated fracture risk and inform treatment decisions, rather than to establish a diagnosis of osteoporosis. Therefore, BMD can serve a diagnostic role (via T-score) or a prognostic role (within risk prediction tools) in a potential screening pathway to prevent fragility fractures.

Treatment

When osteoporosis is diagnosed it can be managed by a combination of lifestyle measures (weight control, exercise, smoking and alcohol avoidance, and calcium/vitamin D supplementation) with pharmacological therapy (43). First-line treatments are usually antiresorptive agents such as bisphosphonates (alendronic acid or risedronate sodium), zoledronic acid or denosumab, which reduce bone turnover and fracture risk, while anabolic or dual-action therapies (for example, teriparatide, abaloparatide, romosozumab) stimulate new bone formation and are reserved for severe or high-risk cases (44). In post-menopausal women oestrogen or selective oestrogen receptor modulators (SERMs), can be used (45). Both classes increase BMD and reduce fractures, though bone-forming therapies are generally more effective in severe disease and are often followed by antiresorptives to sustain benefits (44). Long-term, individualised treatment strategies may include switching or combining agents if side effects or treatment failure occur (44). Both the National Institute for Health and Care Excellence (NICE) and NOGG (18, 22, 46) recommend a comprehensive strategy that includes fracture risk assessment, lifestyle modifications, and pharmacological intervention based on risk level. Bisphosphonates are the foundation of treatment, with newer anabolic agents like romosozumab and abaloparatide reserved for those at very high risk or who cannot use first-line agents. Regular reassessment of fracture risk and attention to adherence ensure long-term effectiveness.

Current policy context and previous reviews

The UK NSC currently recommends against screening for this condition following two reviews of screening for osteoporosis in postmenopausal women in 2013 and 2019 (47). The 2019 review carried out by Solutions for Public Health (48) was an evidence summary review to assess whether systematic population screening for osteoporosis in postmenopausal women should be introduced (48). This review revisited and updated the conclusions of an earlier UK NSC external review published in 2013, which had found the available evidence insufficient to support implementation of a national screening programme (48). The 2019 review covered evidence published between 2011 and 2018, and addressed four key questions: How accurate are screening tests? How effective are the interventions? Do randomised controlled trials demonstrate a clinical benefit of screening? Do UK-based evaluations show screening to be cost-effective? For screening accuracy, the focus was on strategies in which clinical risk assessment tools (such as FRAX) were used as the initial screening step with DXA serving as the reference or confirmatory standard for osteoporosis diagnosis.

The 2019 review incorporated new evidence from 2 randomised controlled trials (RCTs), SCOOP and ROSE (49, 50), since the previous review in 2013 (51), which directly compared screening with usual care. Results showed no significant reduction in osteoporotic fractures among systematically screened women compared to those receiving usual care. Although a variety of screening approaches were identified in the wider literature, few met the inclusion criteria, and all of the review's key questions (the accuracy of screening tests, the effectiveness of interventions, whether RCTs demonstrate a clinical benefit, and whether UK-based evaluations show cost-effectiveness) could not be fully addressed by the evidence identified (48).

Overall, the report concluded that important uncertainties and evidence gaps remained. On this basis, the UK NSC maintained its recommendation against introducing a national population osteoporosis screening programme for postmenopausal women.

In 2024, the UK NSC commissioned an evidence map (52) to assess the volume and type of evidence on key issues in osteoporosis screening since the 2019 review (48). The map addressed 3 questions: what is (1a) the volume and type of evidence exploring the clinical benefit

of whole population screening in women over 65, in reducing osteoporotic fractures compared to standard care? (1b) the volume and type of evidence exploring the clinical benefit of targeted screening of women aged under 65 with risk factors for osteoporosis, in reducing osteoporotic fractures compared to standard care? and (2) the volume and type of evidence exploring the cost-effectiveness of screening (whole population and targeted) for osteoporosis?

Nineteen articles were included: 3 RCTs (SCOOP (50), ROSE (49), SALT/SOS (53)) 1 nonconcurrent cohort study (54), 6 reviews (55-60), 5 sibling papers related to SCOOP (61-65) and 4 related to ROSE (66-69). The RCTs and cohort study examined whole-population screening in women over 65, while all 6 reviews also addressed this question to some degree. However, conclusions as to whether screening was significantly beneficial and the recommendations whether to implement screening were reported to be mixed across studies.

The 2025 evidence map found no evidence on the clinical benefit of targeted screening of women aged under 65 with risk factors for osteoporosis and reported mixed results from reviews on osteoporosis screening in women over 65 (55-60). Although the 6 reviews had similar criteria and used largely the same primary trials, their conclusions differed: some reported small but meaningful reductions in fractures, while others found little-to-no benefit. To understand these disparities and identify evidence gaps relevant to policy, the UK NSC has commissioned the current review.

Objectives

The review consisted of 2 components:

Evidence map

Conduct an updated literature search to identify any new primary studies or reviews reporting on the efficacy or cost-effectiveness of population and targeted screening for osteoporosis in women published since the 2025 evidence map.

Umbrella review

Conduct an umbrella review to examine the included primary studies and reviews (identified in both the previous UK NSC reviews and the update (Q1)) in detail and investigate what factors (methodological or otherwise) could account for differences in conclusions between the reviews.

Evidence map

Aims of the evidence map

Evidence maps are rapid evidence products which aim to gauge the volume and type of evidence relating to a specific topic.

This evidence map has been developed to assess whether there has been any new evidence since a previous review was conducted in November 2024 to evaluate the volume and type of evidence on key issues related to screening for osteoporosis. Any new reviews identified will be included in the umbrella review section of this project.

The aim was to address the following questions:

- Question 1a: What is the volume and type of evidence exploring the clinical benefit of whole population screening in women over 65, in reducing osteoporotic fractures compared to standard of care?
- Question 1b: What is the volume and type of evidence exploring the clinical benefit of targeted screening in women aged under 65 with risk factors for osteoporosis, in reducing osteoporotic fractures compared to standard of care?
- Question 2: What is the volume of evidence exploring the cost-effectiveness of screening (whole population and targeted) for osteoporosis?

The findings of this evidence map will provide data for the umbrella review section of this project.

The aim of this document is to present the information necessary to inform UK NSC decision-making processes.

Search methods and results

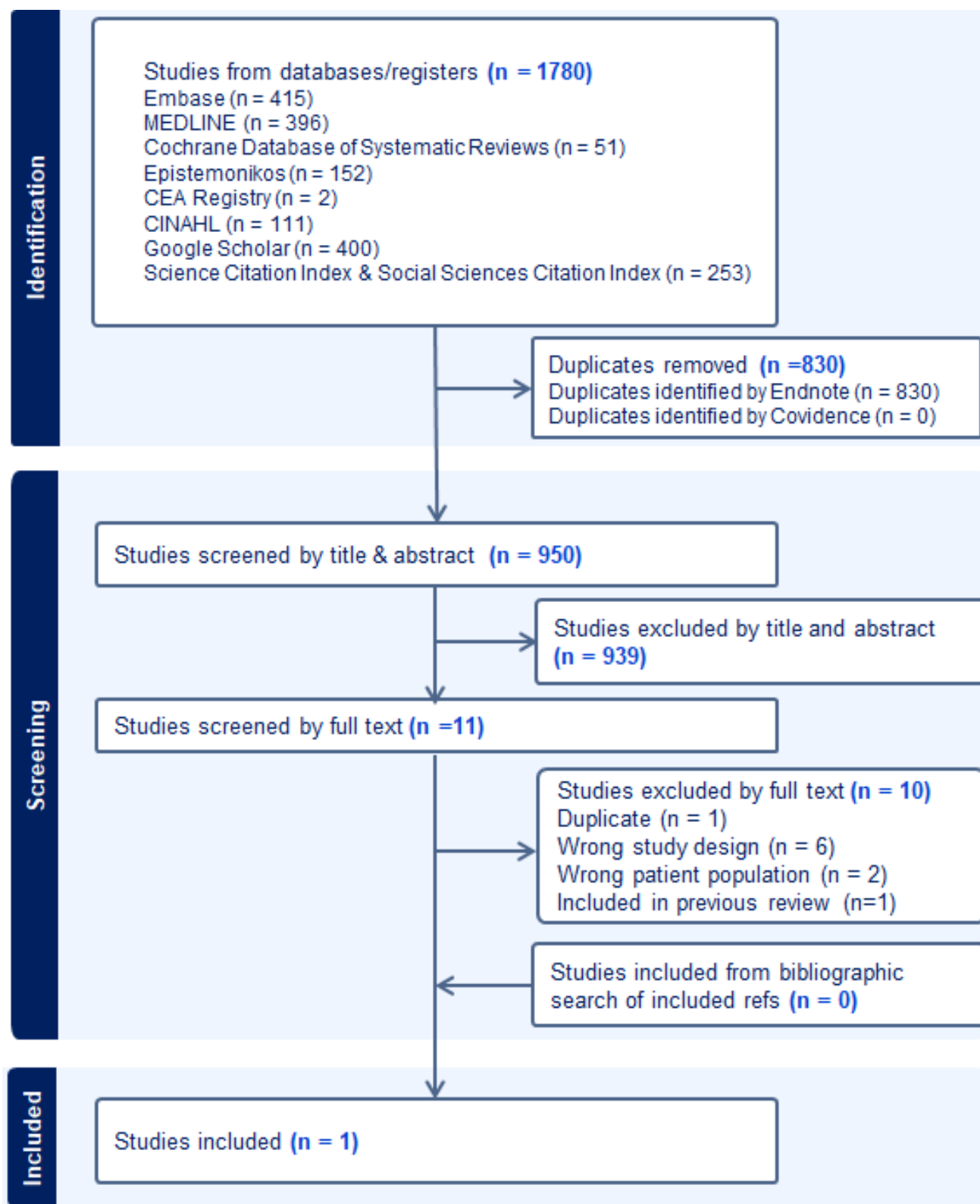
The search strategies from the 2025 evidence map were rerun to find studies published since the last search date (12th November 2024). A date restriction from the start of November 2024 was added to the search for this purpose. The searches were conducted in October and November 2025 using 9 databases: Medline, Embase, CINAHL, Science Citation Index (SCI) and Social Sciences Citation Index (SSCI) (via Web of Science) and the Cochrane Library (Reviews) (accessed 28th October), Epistemonikos, CEA Registry (accessed 29th October) and Google Scholar (searched via Publish or Perish, limited to the first 200 records on the 11th November 2025). The searches were conducted in 9 of the 11 databases used in the previous review; the International HTA Database and Dissertations and Theses Global were excluded because grey literature searching is not included in evidence map methodology. Automatic de-duplication was conducted using Endnote v21.4. The detailed search strategies, including exclusion and inclusion criteria, are available in Appendix 1. Each title and abstract was screened once for eligibility using the Covidence platform (70). A formal quality appraisal of the evidence was not required, given the remit of the evidence map.

Abstract reporting tables are available in Appendix 2.

The search returned 1780 results. After automatic and manual de-duplication, 950 unique references were sifted for relevance to the questions. After screening by title and abstract 11

publications remained, and the full texts were accessed to confirm their eligibility and only one reference, a systematic review (71), was included in the final evidence map. A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram summarising the number of studies included and excluded is presented in Figure 1.

Figure 1. PRISMA flow diagram of included and excluded records.



Summary of findings

Question 1a: What is the volume and type of evidence exploring the clinical benefit of whole population screening in women over 65, in reducing osteoporotic fractures compared to standard care?

One study satisfied the inclusion criteria for this question (71).

The systematic review (SR) by Kahwati 2025 aimed to review evidence on osteoporosis screening in women 65 years or older capturing 3 SRs and 3 RCTs (SCOOP, ROSE and SOS) (71). The review conducted a meta-analysis using data from all 3 RCTs. Of these, SCOOP, and ROSE reported on a screening approach of an initial FRAX assessment followed by BMD testing where indicated, whereas the SOS trial utilised an approach where all women received FRAX and BMD testing plus additional tests. Based on the pooled values, screening was associated with reduced hip (pooled relative risk [RR], 0.83 [95% CI, 0.73-0.93]) and major osteoporotic fractures (pooled RR, 0.94 [95% CI, 0.88-0.99]) when compared to usual care, with absolute risk differences reported as 5 and 6 fewer fractures per 1000 participants screened, respectively.

Question 1b: What is the volume and type of evidence exploring the clinical benefit of targeted screening of women aged under 65 with risk factors for osteoporosis, in reducing osteoporotic fractures compared to standard care?

No new studies met the inclusion criteria for this question.

Question 2: What is the volume of evidence exploring the cost-effectiveness of screening (whole population and targeted) for osteoporosis?

No studies met the inclusion criteria for this question.

Conclusion

Only one study (a review) was identified in the evidence map and that will be included in the umbrella review part of the project (71).

Umbrella review

Aims of the umbrella review

The aim of this umbrella review was to examine the reviews identified in the UK NSC evidence map and their included primary studies in detail, and investigate what factors (methodological or otherwise) could account for differences in conclusions across the reviews.

Decision question

What methodological factors could explain the mixed conclusions across existing reviews on the effectiveness of screening for osteoporosis in women, considering:

- women aged over 65 (whole population)
- women with identifiable risk factors (targeted screening)?²

Methods

Eligibility for inclusion in the umbrella review

Any review publications included in the UK NSC 2025 evidence map and its update.

Data extraction

Data extraction was performed independently by 2 reviewers. Data was collected in Excel; however, only key information necessary to support the findings is presented in tables in the report. The full Excel dataset is provided as a supplementary file to ensure transparency and completeness (Appendix 3).

Data were extracted at both the review and primary study level. For each review³, information was collected on review design and purpose, search methods, eligibility criteria, statistical methods, included studies, quantitative results (including meta-analyses and trial-level data where reported), and author-reported conclusions and methodological limitations. For primary trials included within the reviews, data were extracted on study design, eligibility criteria, interventions and comparators, outcomes and outcome definitions, and reported effect estimates, restricted to those contributing to the review analyses. Full details of the data extraction framework are provided in the protocol.

Methods of analysis/synthesis

We assessed and compared the included reviews and relevant primary studies using a structured framework focusing on four key domains:

- Domain 1: Scope of reviews

² Whilst the evidence map considered targeted screening in women aged under 65, an age limited was not applied in the umbrella review

³ Where publications consider multiple research questions, extraction was limited to only those relevant to the decision question

- Identify and compare the stated aims of each review to determine alignment or divergence in scope.
- Domain 2: Included evidence base
 - Compared reviews in terms of search dates, search strategies, and inclusion criteria, and assessed how these factors contributed to variations in the set of primary studies included.
 - Explored any differences across the primary trials included in the meta-analyses in terms of study designs, populations, and outcome data reported.
- Domain 3: Meta-analyses/numerical data
 - Compared the statistical approaches (for example, effect measures, models, software) used in the meta-analyses.
 - Cross-checked the data inputs from each review (for the meta-analysis or as presented in reviews where no meta-analysis is conducted) against the original trial publications to validate them. Explored any differences between trial data inputs across the reviews.
 - Sense checked the results of each meta-analysis based on data input.
 - Compared the meta-analysis results across studies for each outcome and highlighted any instances where there are key differences such as opposite direction of treatment effects and differences in statistical significance (i.e. the 95% CI does or does not include the null value).
- Domain 4: Author conclusions/interpretation
 - Validated the author reported conclusions of the meta-analysis results. In cases where a review has not conducted a meta-analysis, but reports trial level estimates, these were validated against the reported author conclusions.

Using this framework, differences across domains were synthesised using a structured comparative narrative approach. We compared reviews domain-by-domain using standardised tables of scope (population, intervention, comparator, outcomes and study design [PICOS]), included trials, and numerical inputs and outputs, and then integrated findings across domains to identify the main drivers of differences in results and conclusions.

Appraisal for quality/risk of bias tool

The methodological quality and risk of bias of included studies were assessed using established tools appropriate to study design. Systematic reviews were appraised using ROBIS (72). Randomised controlled trials were assessed using the Cochrane Risk of Bias 2 tool (73) and non-randomised studies using ROBINS-I (74).

Assessments were conducted by one reviewer and checked by a second, with disagreements resolved by consensus or by consultation with a third reviewer where necessary. Risk-of-bias judgements reported in the included reviews were compared with our independent assessments, and any major discrepancies (for example, differences in overall risk-of-bias

judgement) were identified and examined. Where discrepancies were observed, likely reasons such as differences in tools or versions used, signalling questions applied, or availability of information were documented.

Results

Included reviews

A total of 6 relevant reviews were included in the assessment of the reviews (55-59, 71). Five were identified from the 2025 osteoporosis evidence map (55-59). An additional review was identified in the updated 2025 osteoporosis map; Kahwati 2015 (71), this review is an update of a review previously identified in the 2025 evidence map and thus supersedes it (60).

Overview of reviews

A summary of the evidence synthesis product, objectives, and author reported conclusions of each review identified in the osteoporosis evidence map is provided in Table 1.

The reviews were 3 SRs, one rapid relative effectiveness assessment, one synthesis report, and a structured scoping review. They were published between 2003 and 2025. Although most reviews framed their populations broadly, covering adults or older adults in general, the included studies largely involved women, particularly postmenopausal or older women.

The reviews report mixed conclusions on the clinical effectiveness of osteoporosis screening, with one review reporting no clear benefit (56), three identifying small or probable reductions in fracture risk in selected populations (6, 57, 71), and a single review concluding a benefit in older women (59). All reviews included studies relevant to our decision question, namely RCTs that evaluated a combination of risk assessment and DXA compared with a no screening comparator in women aged over 65 years and/or women with identifiable risk factors. Some reviews also included studies that would not have met our evidence map eligibility criteria; and primary study relevance is examined in detail later.

Based on the author reported conclusions of the reviews we have summarised the interpretation of them into the following categories:

- screening not shown to be effective, n=2
- screening potentially effective, n=1
- screening probably effective, n=2
- screening effective, n=1

These mixed conclusions form the basis for our assessment of what factors might explain the variation across reviews, considering both population-level screening in women aged over 65 and targeted screening in women with identifiable risk factors.

Removal of a single review from our assessment

The Health Evidence Network (HEN) synthesis report aimed to assess the effectiveness of osteoporosis prevention and screening (6). The review methods were not fully reported, and it found no direct evidence that screening reduces fracture incidence. As the earliest review in the dataset, with searches extending only to 2003 (with selected studies published in 2004 and 2005 added later), it precedes all major primary trials later captured by subsequent reviews. The

remaining reviews were published between 2019 and 2025, highlighting the considerable gap between this early HEN synthesis and the more recent evidence base. Including this review could confound interpretation of why conclusions differ across reviews, as any discrepancies would likely reflect the historical absence of key primary evidence rather than genuine methodological differences. For this reason, the HEN review is not considered further in our assessment.

Table 1: Summary of reviews

Review/Sponsor	Evidence synthesis product	Objective(s)	Author reported conclusions (Bold text indicates current review interpretation of author conclusions)
Auais 2023 (55) Sponsor NR	Structured scoping review	The primary objective was to review the evidence on how effective clinical fracture-risk assessment tools are in reducing osteoporotic fractures and/or influencing other health outcomes.	Using clinical fracture-risk assessment screening tools was not more effective than usual care in preventing all osteoporosis-related fractures. However, using those screening tools positively influenced women's perspectives on osteoporosis, may have reduced hip fracture risk, and could potentially be cost-effective. SCREENING NOT EFFECTIVE
Gates 2023 (57) Public Health Agency of Canada	Systematic review with meta-analysis	Questions 1a and 1b are relevant for this comparison and described below: 1a: What are the benefits and harms of screening compared with no screening to prevent fragility fractures and related morbidity and mortality in primary care for adults ≥40 years? 1b: Does the effectiveness of screening to prevent fragility fractures vary by screening programme type (i.e., 1-step vs 2-step) or risk assessment tool?	Screening in primary care with clinical FRAX, followed by BMD assessment for increased risk, among selected females aged 65 years and older probably leads to a small reduction in the risk of clinical fragility fracture and hip fracture compared to no screening. Limited direct evidence on harms of screening was available. SCREENING PROBABLY EFFECTIVE
Johnell 2006 (6)* World Health Organisation	Synthesis report	The primary objective was to determine the effectiveness of osteoporosis prevention and screening.	The review did not find direct evidence for fracture reduction as a result of screening for osteoporosis but found good indirect evidence in two specific situations. First, screening is effective for identifying

Review/Sponsor	Evidence synthesis product	Objective(s)	Author reported conclusions (Bold text indicates current review interpretation of author conclusions)
	(Health Evidence Network)		postmenopausal women with low bone mineral density. Second, treating osteoporosis can reduce the risk of wrist and spine fractures in this population. Moreover, there is insufficient evidence relating to the effectiveness of treating low-risk populations. SCREENING POTENTIALLY EFFECTIVE (INDIRECT EVIDENCE ONLY)
Merlijn 2020 (59) Sponsor NR	Systematic review and meta-analysis	To perform a systematic review and meta-analysis to determine the effectiveness of population screening for fracture prevention.	This meta-analysis showed that population screening is effective to reduce osteoporotic fractures and hip fractures. Implementation of screening in older women should be considered as a serious option to prevent osteoporotic fractures, especially hip fractures. SCREENING EFFECTIVE
Kahwati 2025 (71) Agency for Healthcare Research and Quality (AHRQ), US Department of Health and Human Services	Systematic review and meta-analysis	The review considered five research questions. Questions 1 and 3 were relevant to this comparison and are described below: Q1. Does screening for fracture risk or osteoporosis reduce fractures and fracture-related morbidity and mortality in adults?	Screening in higher-risk women 65 years or older was associated with a small absolute risk reduction in hip and major fractures compared with usual care. SCREENING PROBABLY EFFECTIVE

Review/Sponsor	Evidence synthesis product	Objective(s)	Author reported conclusions (Bold text indicates current review interpretation of author conclusions)
		Q3. What are the harms of screening for fracture risk or osteoporosis?	
EUnetHTA OTCA19 Assessment Team 2019 (56) European Commission	Rapid Relative Effectiveness Assessment	Does screening for osteoporosis in the general population offer an advantage over no screening with regard to patient relevant outcomes?	The main finding of this Rapid REA is that screening for osteoporosis in postmenopausal women offers no (or only a small) advantage with regard to symptomatic fractures. Since the available moderate quality studies do not show screening to have any effect on the incidence of symptomatic fractures, there is probably little or no benefit to screening for osteoporosis in postmenopausal women SCREENING NOT EFFECTIVE

Abbreviations: BMD, bone mineral density; FRAX, Fracture Risk Assessment Tool; NR, not reported; REA, Rapid Evidence Assessment.

*The HEN synthesis report is in grey because this review was excluded from the full assessment of the reviews.

Assessment of reviews

Results are presented by domain. Although multiple methodological and scope differences were identified across reviews, we examine whether these differences materially influenced the numerical results or the conclusions drawn.

Domain 1: Scope of reviews

Overall, the reviews demonstrated broad alignment in their objectives and PICOS frameworks, each assessing screening strategies combining clinical risk assessment and/or BMD testing against no screening or usual care. Differences were mainly observed in the inclusion of additional study types beyond RCTs, the screening interventions and in the range of outcomes considered. The following subsections describe the scope of the reviews in more detail according to the review objectives and each PICOS element.

Review objectives

Review objectives were similar across the reviews in all seeking to evaluate the effectiveness of screening or risk assessment strategies for osteoporosis or fracture prevention in adults (Table 1).

Population

Two reviews considered adults ≥ 40 years (57, 71) whereas all other reviews considered a general asymptomatic population.

Interventions

All reviews evaluated population-based screening strategies designed to identify individuals with osteoporosis or with increased risk of fragility fractures, but they differed in how the intervention was defined. Most reviews conceptualised screening as a multistep process combining a clinical fracture-risk assessment tool, typically FRAX, with BMD measurement by DXA where indicated, followed by pharmacological treatment for those classified as high risk.

However, important notable differences included:

- Gates 2023 permitted only fracture risk assessment (validated or non-validated tools), only BMD measurement with DXA, and would also permit vertebral fracture assessment (VFA) or spinal radiography in addition to DXA or risk assessment and DXA (57).
- Kahwati 2025 (71) and Merlijn 2020 (59) applied similar definitions, focusing on risk-assessment and/or DXA-based screening linked to subsequent treatment.
- EUnetHTA 2019 used a broader definition that also allowed radiographic assessment of vertebral fractures as a screening method (56).
- Auais 2023 specifically restricted inclusion to screening using validated clinical risk tools (FRAX, GARVAN, CAROC) regardless of whether BMD was measured (55).

These definitional differences, particularly regarding whether DXA measurement was required, whether radiography was accepted, and whether validated risk tools were mandatory represent a key methodological distinction which contributed to variation across reviews in the studies included.

Comparator

The reviews compared screening interventions with no screening or usual care. In a single review, whilst a comparator was not specified based on the studies included, comparisons were made between different screening strategies (e.g. BMD versus risk assessment tools) or alternative risk assessment tools (e.g. Simple Calculated Osteoporosis Risk Estimation (SCORE) vs the Study of Osteoporotic Fractures (SOF) index) in addition to no screening (55).

Outcomes

All reviews assessed fracture-related outcomes as the primary measure of screening effectiveness, typically including osteoporotic, clinical, fragility, or hip fractures. Several also evaluated secondary outcomes such as mortality, health-related quality of life, pain, and functional status. A subset of reviews considered harms of screening, including adverse events, overdiagnosis, and overtreatment, as well as healthcare resource use. Overall, while the emphasis on fracture outcomes was consistent, the breadth of secondary outcomes and harms varied across reviews.

Study design

Randomised controlled trials were the principal evidence source across all reviews. Gates 2023 (57) supplemented these with other clinical trials, whereas Kahwati 2025 also considered SRs (71).

Domain 2: Included evidence base

Searches

All reviews conducted literature searches in the major biomedical databases, most commonly MEDLINE, Embase, and the Cochrane Library. All supplemented these searches by consulting clinical trial registries (such as ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform) and scanning reference lists of included studies or relevant SRs. The EUnetHTA OTCA19 review also searched guideline and health technology assessment (HTA) agency websites (e.g. NICE, AHRQ) to identify additional grey literature. Overall, the search strategies were comprehensive and broadly similar across reviews, suggesting that differences in search sources are unlikely to have materially affected the studies identified.

The searches across reviews covered broadly overlapping but not identical time periods. The earliest review searched for studies published up to 2019, while the most recent extended coverage to 2024.

Summary of evidence base in each review

The primary studies included across the reviews are summarised in Table 2 and Table 3. Of the trials included across the reviews, 3 pivotal RCTs are relevant to the evidence map (SOS/ROSE/SCOOP), as they evaluate screening strategies involving risk assessment and DXA, include a no-screening comparator, and were conducted

in women aged ≥ 65 years. Nevertheless, the reviews differed in scope and there were variabilities in the included studies. The main drivers for these differences are:

1. Comparator definitions. All reviews required a no-screening or usual-care control and therefore excluded studies comparing alternative screening strategies except for Auais 2023 which did not state a restriction on comparator, and this is why this review included OPRA (which does not include a no screening comparator arm) (55).
2. Search dates. The search date determined whether the most recently published trials were included. EUnetHTA 2019 (56) did not include SOS due to its publication date. Notably, despite being published prior to the SOS paper, Merlijn 2020 (59) included data from SOS, as one of the review's authors was also an author on the SOS publication and therefore had access to the SOS data.
3. Study designs: All reviews restricted inclusion to primary trials that were RCTs except Gates 2023 which included any primary study design (57). This is why this review included Kern 2005 (a non-concurrent cohort study) (54).
4. Screening pathways: The reviews differed in how screening pathways were defined but most studies included in the reviews included pathways relevant to the evidence map (any combination of risk assessment and DXA). Exceptions included:
 - COSHIBA: The screening intervention included spinal X-ray (no BMD and no validated risk tool) as the objectives of the study were to identify postmenopausal women with osteoporotic vertebrae fractures and increase appropriate management of osteoporosis (75). COSHIBA was only relevant for inclusion in EUnetHTA 2019 (56), which used a broader definition of screening that allowed standard radiographs to detect vertebral fractures.
 - APOSS: APOSS falls outside the scope of most reviews because the comparator and treatment pathway do not match the decision questions (76). It used DXA alone without risk assessment and recommended treatment of high-risk participants with hormone replacement therapy which authors of some of the reviews confirmed is no longer used for the treatment of osteoporosis. As a result, it was excluded from all but Gates 2023 (57), and even in this review its results were not incorporated into the meta-analysis due to uncertainty in the evidence.

In summary, although there was some variability in the studies included across reviews, they largely drew on the same underlying evidence base, centred on the 3 pivotal RCTs. In most cases, meta-analyses did not incorporate additional studies beyond these trials. For this reason, the summary of primary studies in the next section focuses on the pivotal RCTs.

Table 2: Summary of included primary studies in the reviews

Study (year published, design)	Relevant to NSC evidence map decision question?	Merlijn 2020 (59), n=3	Gates 2023 (57), n=5	Kahwati 2025 (71), n=3	EUnetHTA 2019 (56), n=3*	Auais 2023 (55), n=4 [†]
ROSE (2018, RCT) (49)	✓	✓	✓	✓	✓	✓
SCOOP (2018, RCT) (50)	✓	✓	✓	✓	✓	✓
SOS (2019, RCT) (53)	✓	✓	✓	✓	X [publication date]	✓
OPRA (2005, RCT) (77)	X (no standard care/no screening comparator)	X (comparator)	X [‡] (comparator)	X (comparator)	X (comparator)	✓
COSHIBA (2012, RCT) (75)	X (Age 45-54/no risk factors years & intervention doesn't include BMD)	X (no BMD)	X (no BMD)	X (no BMD or FRAX)	✓	X (not validated risk tool)
APOSS (1997, RCT) (76)	X (Intervention includes BMD only)	X (HRT trt)	✓ (but excluded from meta-analyses)	X (poor quality & HRT no longer standard practice in US)	X (HRT trt)	X (HRT trt)

Kern 2005 (2005, non-random allocation) (54)	X (Intervention includes BMD only)	X (non-RCT)	✓ (included in meta-analysis of hip fractures & mortality only)	X (non-RCT)	X (non-RCT)	X (non-RCT)
--	------------------------------------	-------------	---	-------------	-------------	-------------

Abbreviations: BMD, bone mass density; FRAX, Fracture Risk Assessment Tool; HRT, hormone replacement therapy; RCT, randomised controlled trial; trt, treatment.

Rationale for study relevant to the UK NSC evidence map decision question or its eligibility for inclusion into reviews provided in brackets.

Studies in grey were not relevant to the UK NSC evidence map decision question.

*Core study pool of 3 conventional RCTs included and considered here. The review also included 5 enrichment design RCTs which provide supplementary evidence on treatment efficacy, not on screening effectiveness and therefore are not considered here.

†only 4/6 studies included reported outcomes relevant to the current review decision question. The other 2 qualitative studies included reported perspectives, experiences and acceptance of screening and are not considered here.

‡This study was included for KQ1b (Does the effectiveness of screening to prevent fragility fractures vary by screening program type (i.e., 1-step vs. 2-step) or risk assessment tool) which was not considered in the current analysis

Table 3: Summary of primary studies included in the reviews

Study, country, inclusion	Population characteristics	Intervention/comparator	Primary outcome, analysis populations, follow-up
ROSE (49) RCT Denmark Women aged 65–80 years	<u>Screening (n=17,072; ITT)</u> Age, median (IQR): 71 (68–76) years 10-year MOF risk (FRAX), median (IQR): 20% (15–27) 10-year hip fracture risk (FRAX), median (IQR): 6.7% (3.9–11) <u>Control (n=17,157; ITT)</u> Age, median (IQR): 71 (68–76) years 10-year MOF risk (FRAX), median (IQR): 20% (15–27) 10-year hip fracture risk (FRAX), median (IQR): 6.7% (3.9–11)	All patients completed FRAX-Denmark Screening: DXA when 10-year risk of MOF was $\geq 15\%$. Treatment as per recommendations based on the Danish National guidelines Control: Women not informed of baseline FRAX risk	Primary outcome: ≥ 1 MOF (records - ICD-10 codes) Analysis populations: ITT (all women randomised); PP-1 (women who returned the questionnaire with adequate information to calculate a FRAX value); PP-2 (women with $FRAX \geq 15\%$) Follow-up: median 5 (Q1 4.6, Q3 5.3) years
SCOOP (50) RCT England Women aged 70–85	<u>Screening (n=6,233)</u> Age, mean (SD): 75.4 (4.16) years 10-year MOF risk (FRAX), mean (SD): 19.3% (8.9) 10-year hip fracture risk (FRAX), mean (SD): 8.5% (7.4) <u>Control (n=6,250)</u> Age, mean (SD): 75.5 (4.14) years 10-year MOF risk (FRAX), mean (SD): 19.3% (8.8) 10-year hip fracture risk (FRAX), mean (SD): 8.5% (7.3)	All patients completed FRAX Screening: DXA when 10-year risk of hip fracture was $>$ threshold probability for their age. Then the 10-year hip fracture risk was recalculated with the BMD and a final risk category communicated to the participants GP (to receive standard care).	Primary outcome: ≥ 1 osteoporotic related fracture (self-reported and verified) Analysis population: ITT (participants analysed according to the group they were randomly assigned whether screening was completed; baseline questionnaire must have been completed) Follow-up: 5 years

Study, country, inclusion	Population characteristics	Intervention/comparator	Primary outcome, analysis populations, follow-up
		Control: Women not informed of baseline FRAX risk	
SOS (53) RCT Netherlands Women aged 65 to 90 years; with ≥ 1 clinical risk factor for fractures*, as assessed with a baseline questionnaire	<u>Screening (n=5,575)</u> Age, mean (SD): 75 (6.7) years 10-year MOF risk (FRAX), mean (SD): 24.6% (10.8) 10-year hip fracture risk (FRAX), mean (SD): 11.6% (10.5) <u>Control (n=6,250)</u> Age, mean (SD): 75 (6.7) years 10-year MOF risk (FRAX), mean (SD): 24.3% (10.5) 10-year hip fracture risk (FRAX), mean (SD): 11.3% (10.2)	Screening: DXA and VFA. The 10-year risk of MOF (FRAX-UK) above age-dependent thresholds in combination with DXA-T score or prevalent fracture was an indication for treatment with anti-osteoporosis medication Control: At baseline, those with an indication for DXA and VFA, according to the Dutch GP guideline, were notified accordingly and were advised to contact their GP as part of usual care. All participants offered the same screening intervention after study completion	Primary outcome: any type of fractures (self-reported and verified) Analysis population: ITT (excluded patients lost to follow-up) Follow-up: mean 3.7 years
COSHIBA (75) RCT England	<u>Screening (n=1,062)</u> Age, mean (SD): 72.7 (4.3) years <u>Control (n=2,138)</u> Age, mean (SD): 72.6 (4.3) years	Screening: Data collected on four clinical risk factors and if the score was below 4 (high risk of prevalent	Primary outcome: Self-reported new prescription of medications for treatment osteoporosis at the 6-month follow-up

Study, country, inclusion	Population characteristics	Intervention/comparator	Primary outcome, analysis populations, follow-up
Women aged 65 to 80 years		VF), offered a thoracolumbar radiograph. Usual NHS mechanisms for reporting results used to mimic real-life situation Control: All participants received a “healthy bones” leaflet giving information on healthy lifestyles	Analysis population: ITT (all secondary fracture outcomes self-reported) Follow-up: 12 months
Kern 2005 (54) Non-concurrent cohort study US Age- and sex-stratified random sampling of Medicare eligibility lists	<u>Screening (n=1,422)</u> Age, mean (SD): 76.3 (4.8) years <u>Control (n=1,685)</u> Age, mean (SD): 72.6 (4.9) years	Screening: DXA scans offered and results sent to the participants primary care provider Control: Usual care	Primary outcome: Hip fractures (active and passive surveillance) Analysis population: Analytic cohort Follow-up: Mean 4.9 years
APOSS (76) RCT Scotland Women aged between 45 and 54 years of age	<u>Screening (n=2,400)</u> Age, mean (SD): 58.4 (4.0) years <u>Control (n=2,400)</u> Age, mean (SD): 58.4 (3.3) years	Screening: Participants invited for DXA, for those considered ‘at risk’ the GP was advised to offer HRT when the women reached the menopause Control: Women were not contacted	Primary outcome: Drug use Analysis population: ITT (all subject who responded to the follow-up questionnaire) and PP (only screened women who attended screening) [fracture outcomes self-reported and verified] Follow-up: 9 years

Study, country, inclusion	Population characteristics	Intervention/comparator	Primary outcome, analysis populations, follow-up
OPRA (77) RCT US Women aged 60–80 not taking hormone therapy or osteoporosis medication	<u>Universal testing (n=415)</u> <i>Age, mean (SD): 69.1 (5.5) years</i> <u>SCORE testing (n=576)</u> <i>Age, mean (SD): 69.3 (5.5) years</i> <u>SOF-Testing (n=2,176)</u> <i>Age, mean (SD): 70.3 (5.5) years</i>	Screening programme identical for all 3 arms but differed in how liberally BMD testing was offered Universal testing: All women offered BMD testing SCORE testing: Women asked to complete SCORE. Those with a high-risk score were invited for BMD testing. SOF-Testing: Women asked to complete SOF. Those with 5 or more risk factors were invited for BMD testing.	Primary outcome: Initiation of osteoporosis therapy Analysis population: ITT (all subject invited to participate) and participant-only analysis (subgroup that responses and became intervention participants) [fracture outcomes verified with ICD-9 codes] Follow-up: mean 28 (range 24-33) months

Abbreviations: BMD, bone mass density; DXA, Dual-energy X-ray absorptiometry; GP, general practitioner; HRT, hormonal replacement therapy; ICD, International Classification of Diseases; IQR, interquartile range; ITT, intention to treat; MOF, major osteoporotic fracture; PP, per-protocol; RCT, randomised controlled trial; SCORE, Simple Calculated Osteoporosis Risk Estimation; SD, standard deviation; SOF, Study of Osteoporotic Fracture; VFA, vertebral fracture assessment.

Studies in grey were not relevant to the UK NSC evidence map decision question

* a previous fracture after age 50 years, a parental hip fracture, low body weight, rheumatoid arthritis, early menopause, malabsorption syndrome, chronic liver disease, type I diabetes mellitus, or immobility

Summary of pivotal RCTs

All 3 pivotal RCTs compared screening strategies with usual care. ROSE and SCOOP evaluated population screening strategies, using an initial FRAX assessment followed by DXA for women above a predefined FRAX threshold (49, 50). Notably in SCOOP (50) women underwent the baseline questionnaire before randomisation, whereas in ROSE (49) women underwent the questionnaire after randomisation. In contrast, SOS evaluated a targeted screening strategy, enrolling women with at least one clinical risk factor for osteoporosis as identified by a baseline questionnaire and then randomised participants; the screening intervention included DXA, VFA, FRAX and blood tests for all women (53). The higher baseline fracture risk of participants in SOS (53) reflects this enriched population and is evident in the higher baseline FRAX-estimated 10-year risk of major osteoporotic and hip fracture of participants compared with those in ROSE (49) and SCOOP (50) (Table 3).

Usual care served as the comparator in all 3 RCTs, although this was not a true 'no screening' arm. Completion of baseline questionnaires and FRAX assessments likely increased disease awareness across both arms, despite results not being routinely disclosed to participants in the control arms of ROSE and SCOOP, leading to higher-than-expected rates of DXA in the control groups. For example, the authors of ROSE reported that 25% of women in the control group had a DXA scan after the index date versus 48% in the screening group (49). Similarly, 23% of women in the screening group received osteoporosis medication compared with 18% in the control group (49).

All 3 RCTs reported analyses based on intention-to-treat (ITT) populations. ROSE also presented 2 additional per-protocol (PP) populations: PP-1 (women with sufficient information to calculate FRAX) and PP-2 (women with FRAX $\geq 15\%$) which represents a high fracture risk subgroup analysis (49). Generally, no benefit was observed in the ITT analysis, while progressively more favourable effects were seen in the PP-1 and PP-2 populations, consistent with increasing engagement with screening and restriction to higher-risk participants. Although the ROSE PP-2 population (49) was restricted to women with FRAX $\geq 15\%$, the median FRAX risk in this group remained lower than the mean baseline risk in SOS, which enrolled women with at least one clinical risk factor for fracture (53). This reflects the differences in trial design and risk-enrichment strategies.

On the basis of overall baseline risk, the ROSE PP-2 population represents a more enriched subgroup than ROSE ITT or PP-1 (49), and may be considered broadly comparable to SOS ITT (53) in terms of fracture risk, both representing targeted screening although the populations are not directly equivalent. In contrast, ROSE PP-1 (49), which included all women with sufficient data to calculate FRAX, aligns more closely with the SCOOP ITT population for population screening (50).

Despite these differences, reviews undertook meta-analyses and all reviews that included all 3 trials selected the ROSE PP-1 population for meta-analysis (49), with the exception of EUnetHTA 2019 (56) which instead combined ROSE ITT (49) data with SCOOP ITT data (50). Notably, although reviews pooled all 3 trials, ROSE/SCOOP and SOS evaluated different screening models in populations with different baseline risk, which may limit the interpretability of combined estimates. Moreover, across reviews, the apparent overall benefit is driven in part by reliance on the ROSE PP-1 analysis; however, the ROSE ITT analysis is more relevant to a screening context as it reflects an 'offer of screening' to the eligible population, did not demonstrate an effect in favour of screening.

All 3 trials were judged to have some concerns of bias (Appendix 3 figure 1.1). Although all 3 trials were well randomised and had complete follow-up, their pragmatic design precluded

blinding of participants and study personnel, introducing methodological limitations related to intervention adherence, contamination (ad hoc screening and treatment in control arms), and outcome ascertainment (self-reported fractures) are likely to have biased effect estimates towards no effect through dilution of between-group differences, potentially underestimating the true effect of osteoporosis screening. Conversely, recruitment of risk-enriched populations may have inflated effect estimates, partially counterbalancing dilution biases while reducing generalisability.

Time-to-event analyses and their implications for meta-analysis

The 3 pivotal RCTs all used time-to-event methods for analysing fracture outcomes, applying Cox proportional hazards models (SOS and SCOOP) or competing-risk survival analysis (ROSE) (49). Consequently, meta-analyses would typically input the hazard ratios from these models (as log [HR] and SE) rather than replacing them with relative risks calculated from dichotomous event counts. This is because, time-to-event models appropriately account for censoring and variation in follow-up time, whereas risk ratios derived from dichotomous event counts at a fixed time point do not.

Patterns observed across the Kaplan–Meier curves in ROSE and SCOOP further reinforce this point (49, 50). Both trials show that the benefits of screening and targeted treatment accumulate over time, with little separation in the early years but increasingly divergent curves with longer follow-up. In ROSE, the cumulative incidence curves for major osteoporotic fractures, hip fractures and all fractures diverge progressively in the PP-2 population (49). Similarly, in SCOOP, the curves for time to first osteoporotic fracture and hip fracture in the ITT population show widening separation across follow-up (50).

These findings have direct implications for meta-analysis. Because fracture benefits emerge gradually, dichotomous count data fail to capture the time-dependent nature of the effect and ignore censoring. Hazard ratios from time-to-event models therefore provide a more appropriate and comparable basis for pooling evidence across trials with differing follow-up durations and should be preferred for meta-analytic synthesis of fracture outcomes.

Methodological quality of the reviews (ROBIS assessment)

Across the four SRs (Merlijn 2020, Gates 2023, EUnetHTA 2019 and Kahwati 2025,) ROBIS assessments indicated an overall low risk of bias (Appendix 3 figure 1.3), with generally well-defined eligibility criteria, appropriate study designs and populations, and clearly reported methods (56, 57, 59, 71). All four reviews used eligibility criteria that were aligned with their review. Searches were broadly comprehensive in Gates 2023, EUnetHTA 2019 and Kahwati 2025, drawing on multiple databases and trial registries and using additional methods (reference checking, author contact, surveillance searches) (56, 57, 71). In contrast, Merlijn 2020 relied primarily on MEDLINE and Embase, with no reported supplementary search strategies, leading to a high concern judgement for the identification and selection domain (59).

Data collection and study appraisal processes were robust across reviews, typically involving dual-reviewer extraction or checking and formal risk-of-bias assessment with appropriate tools. Synthesis methods were also judged appropriate in all cases, with random-effects meta-analysis used where pooling was justified, low or explored heterogeneity, and sensitivity analyses undertaken to test robustness. Common limitations included restriction to English-language publications and a small number of eligible RCTs, which precluded meaningful use of formal methods to assess publication bias (such as funnel plots or related statistical tests); however, these limitations were considered unlikely to introduce substantial bias into the overall conclusions.

Domain 3: Meta-analyses/numerical data

Auias 2023 provides a narrative summary of the scoping review and did not report any numerical data from the studies it included so it is not considered within this section (55).

Statistical approaches

Four reviews undertook pairwise meta-analyses using either RevMan (Merlijn 2019 (53), statistical method not reported; Gates 2023 (57), DerSimonian and Laird model), SAS (EUnetHTA (56); statistical method not reported) or STATA (Kahwati 2025 (71); DerSimonian and Laird model). Merlin 2019, Gates 2023 and Kahwati 2025 (53, 57, 71) employed random-effects (RE) models, whereas EUnetHTA utilised a fixed-effect (FE) model (56). Given the limited number of studies included in these analyses (≤ 4) and the inability to reliably estimate between-study heterogeneity within RE frameworks, the RE models would have effectively converged toward FE estimates. Consequently, the choice of statistical modelling approach is unlikely to have materially influenced the results.

Trial data considered in the reviews

For mortality and all fracture-related outcomes, all trial-level data reported in the reviews and used as inputs for the meta-analyses were checked for consistency with the corresponding outcomes reported in the original trial publications. Across reviews, the following approaches were used:

- **Merlin 2019 (53), EUnetHTA 2019 (56) and Kahwati 2025 (71)** used trial level dichotomous data for meta-analysis and therefore relied on unadjusted event counts.
- **Gates 2023** used log hazard ratios (log[HR]) and their standard errors, incorporating unadjusted HRs where reported (SCOOP reported only adjusted HRs) (57).

As outlined in Domain 2, time-to-event models (and their HRs) account appropriately for censoring and variation in follow-up time, whereas risk ratios derived from dichotomous data do not. Hazard ratios therefore remain the most robust and methodologically appropriate effect measure for meta-analysis of fracture outcomes

Despite the mixed analytical approaches used across reviews, the pooled effect estimates were nevertheless highly consistent (Table 4). This convergence is likely coincidental and may reflect the broadly similar follow-up durations across trial arms, low event rates, modest true effects, and limited censoring events.

Notably, all reviews drew on the same primary trial publications. Ten-year follow-up data from the ROSE trial was published in May 2024 and, given the timing, most reviews would not have been expected to include it (68). However, although Kahwati 2025 reported active surveillance up to July 2024, this publication was still not incorporated. The ten-year ROSE results were extracted for completeness, but they are consistent with the five-year findings in both the magnitude of effect and statistical significance. Importantly, inclusion of the ten-year ROSE follow-up data in any of the reviews would not have altered the overall results or conclusions, as the findings remain consistent with the five-year outcomes.

Meta-analysis results

In consideration of the relative effect estimates presented across the reviews (meta-analyses or trial level data presented), the conclusions are consistent in terms of direction of effect size and statistical significance (Table 4).

In addition to the relative effect estimates, all reviews presenting numerical data also reported absolute risk differences or absolute risks, except Merlijn 2019. These were generally derived by applying the pooled relative risk to either the median or pooled control-group event rate across the included trials. As Gates 2023 (57) and Kahwati 2025 (71) represent the most recent and relevant reviews, each drawing on at least the same 3 main screening trials, their absolute risk difference estimates are tabulated for comparison (Table 5).

Notably, Gates (2023) also presented absolute risk differences based on general-population baseline risks from Statistics Canada for women aged ≥ 65 years for GRADE assessment of the evidence (57). This is crucial because the magnitude of baseline fracture risk directly determines the size of the absolute benefit: even identical relative effects (e.g., HRs around 0.94–0.95) translate into materially different absolute risk differences in high-risk versus lower-risk populations. Using population-relevant baseline rates ensures that estimated benefits reflect whether screening yields a clinically meaningful reduction in fractures in real-world settings.

In summary, although the underlying trial results were similar across reviews, conclusions varied depending on how the size of the benefit was judged. Some reviews focused on absolute reductions in fracture risk, which depend on assumptions about baseline risk, while others focused on relative effects. These differences in emphasis appear to have played a larger role in shaping review conclusions than any substantive differences in the trial evidence itself.

Table 4: Summary of meta-analyses and trial level estimates considered in the reviews. Estimates are HR or RR (95% CI) for screening intervention versus control unless stated otherwise

Trials	Trial outcome definitions	Merlijn 2020 (53)	Gates2023 (57)	Kahwati 2025 (71)	EUnetHTA 2019 (56)
All fractures					
COSHIBA (75)	Self-reported new fractures (fore-arm, hip, vertebral, other)				OR: 0.60 (0.35-1.03)
SCOOP (50)	Any clinical fractures (included all occurring fractures, including fractures of the hands, feet, ankle, face, and skull)	0.94 (0.86-1.03)		0.95 (0.88-1.03)	0.94 (0.86-1.03)
SOS (53)	All fractures	0.97 (0.87-1.08)		0.97 (0.89-1.08)	
Meta-analysis	All fractures	0.95 (0.89-1.02)		0.96 (0.90-1.02)	
Osteoporotic fractures					
ROSE (49)	All potential osteoporosis fractures (excluding fractures of fingers, toes, skull, or face)	0.97 (0.89-1.06)		0.98 (0.90-1.06)	1.01 (0.95, 1.07)
SCOOP (50)	Osteoporosis-related fractures (all fractures excluding the hands, feet, nose, skull, or cervical vertebrae; vertebral fractures documented within 6 months of randomisation were excluded due to uncertainty about the actual date of occurrence)	0.94 (0.85-1.03)		0.95 (0.87-1.04)	0.94 (0.85-1.03)
SOS (53)	Osteoporotic fractures defined as all fractures except for skull, finger, hand, toe, and foot fractures	0.91 (0.89-1.00)		0.93 (0.83-1.04)	
Meta-analysis	Osteoporotic fractures	0.95 (0.89-1.00)		0.95 (0.91-1.01)	0.99 (0.94 -1.04)
Major osteoporotic fractures					

ROSE (49)	Major osteoporotic fractures (hip, clinical vertebral, wrist or humerus fracture)	0.91 (0.82-1.00)	0.92 (0.83-1.02)	0.93 (0.84-1.02)	0.99 (0.92-1.06)
SCOOP (50)	Osteoporosis-related fractures (all fractures excluding the hands, feet, nose, skull, or cervical vertebrae; vertebral fractures documented within 6 months of randomisation were excluded due to uncertainty about the actual date of occurrence)		0.94 (0.85-1.04)	0.95 (0.87-1.04)	
SOS (53)	Major osteoporotic fractures were defined as all hip, vertebral, wrist, and humerus fractures	0.91 (0.80-1.04)	0.92 (0.81-1.05)	0.93 (0.82-1.05)	
Meta-analysis	Major osteoporotic fractures	0.91 (0.84-0.98)	0.93 (0.87-0.99)	0.94 (0.88-0.99)	
Hip fractures					
COSHIBA (75)	Hip fractures				1.01 (0.25-4.74)
ROSE (49)	Hip fractures	0.83 (0.67-1.01)	0.84 (0.68-1.03)	0.84 (0.69-1.03)	1.01 (0.89-1.14)
SCOOP (50)	Hip fractures (verified fractures with a specific description of neck of femur or proximal femur. Fractures described as subtrochanteric, femoral shaft, distal femur, or simply femoral were not categorised as hip fractures)	0.72 (0.59-0.88)	0.73 (0.60-0.88)	0.75 (0.62-0.92)	0.72 (0.59-0.89)
SOS (53)	Hip fractures	0.91 (0.71-1.15)	0.91 (0.72-1.15)	0.91 (0.72-1.15)	
Kern 2005 (54)	Hip fractures		0.61 (0.35-1.06)		
Meta-analysis	Hip fractures	0.80 (0.71- 0.91)	0.80 (0.71-0.91)	0.83 (0.73-0.93)	
All-cause mortality					
SCOOP (50)	Mortality	1.05 (0.93-1.19)	1.05 (0.94-1.18)	1.05 (0.94-1.18)	1.05 (0.93-1.19)
SOS (53)	Mortality	1.03 (0.91-1.17)	1.02 (0.91-1.15)	1.02 (0.91-1.15)	
ROSE (49)	Mortality			0.97 (0.92-1.03)	
Kern 2005 (54)	Mortality		0.90 (0.78-1.04)		
Meta-analysis		1.04 (0.95-1.14)	1.00 (0.92-1.09)	0.99 (0.95-1.04)	

UK NSC external review – Screening for osteoporosis: An evidence map and umbrella review for the UK National Screening Committee, [Date of review completion]

Abbreviations: CI, confidence interval; HR, hazard ratio; OR, odds ratio; RR, relative risk. Green shading indicates a statistically significant results favouring screening versus no screen. Red shading indicates a non-statistically significant result.

Table 5: Summary of meta-analysis relative treatment effect results and absolute effects. Estimates are for screening intervention versus no intervention.

Outcomes	Gates 2023 (57)			Kahwati 2025 (71)	
	HR (95% CI)	Absolute effect (control rate from trials)	Absolute effect (control rate from general population of women aged ≥65)	RR (95% CI)	Absolute effect (control rate from trials)
All fractures				0.96 (0.90-1.02)	6 fewer per 1000 (14 fewer to 3 more)
Osteoporotic fractures				0.95 (0.91-1.01)	6 fewer per 1000 (11 fewer to 1 more)
Major osteoporotic fractures	0.93 (0.87-0.99)	5.9 fewer per 1000 (10.9 fewer to 0.8 fewer)	11.8 fewer per 1000 (21.8 fewer to 1.7 fewer)	0.94 (0.88-0.99)	6 fewer per 1000 (12 fewer to 1 more)
Hip fractures	0.80 (0.71-0.91)	6.2 fewer per 1000 (9.0 fewer to 2.8 fewer)	4.0 fewer per 1000 (5.8 fewer to 1.8 fewer)	0.83 (0.73-0.93)	5 fewer per 1000 (7 fewer to 2 more)
All-cause mortality	1.00 (0.92-1.09)	No difference per 1000 (7.1 fewer to 5.3 more)	No difference per 1000 (4.6 fewer to 5.1 more)	0.99 (0.95-1.04)	1 fewer per 1000 (5 fewer to 4 more)

Abbreviations: CI, confidence interval; HR, hazard ratio; RR, relative risk.

Domain 4: Author conclusions/interpretation

Our independent validation of author reported conclusions from the reviews based on the trial level numerical data presented or meta-analysis results is presented in Table 6. In most cases, review conclusions were consistent with the underlying evidence. One exception was identified in the interpretation of fracture outcomes in Merlijn 2020, where conclusions referring broadly to reductions in osteoporotic fractures extended beyond the statistically significant findings observed for major osteoporotic fractures and hip fractures. The pooled estimate for all osteoporotic fractures (HR 0.95, 95% CI 0.89–1.00) did not demonstrate a statistically significant reduction, whereas pooled estimates for major osteoporotic fractures and hip fractures did.

Auais 2023 was a structured scoping review that concluded clinical fracture-risk assessment screening tools were not more effective than usual care for preventing osteoporosis-related fractures (55). The review provided a qualitative synthesis only: it summarised the findings of each trial descriptively, discussed statistically significant results, but did not present numerical estimates or indicate the magnitude of relative effects. We checked the review's main findings against the results reported in the primary trial publications and found them to be consistent. Because no meta-analysis was undertaken, the review did not capture that the 3 trials, although non-significant individually, can yield a statistically significant effect when pooled—for example, for major osteoporotic fractures. A similar pattern is seen for hip fractures: SCOOP reported a statistically significant reduction, while the other 2 trials did not, yet the pooled estimate in meta-analyses becomes statistically significant. This likely explains the divergence between the conclusions of Auais 2023 (55) and those of reviews that incorporate meta-analysis.

Table 6: Validation of review author reported conclusions with numerical data and meta-analysis results presented

Review	Author reported conclusions relevant to decision question (Bold text indicates current review interpretation of author conclusions)	Independent validation
Auais 2023 (55)	Using clinical fracture-risk assessment screening tools was not more effective than usual care in preventing all osteoporosis-related fractures. However, using those screening tools positively influenced women's perspectives on osteoporosis, may have reduced hip fracture risk, and could potentially be cost-effective. SCREENING NOT EFFECTIVE	The authors' statements about the effectiveness of FRAX-based screening were consistent with the results reported in the 3 primary trials (although the authors did not present any numerical results). • Their conclusion that screening did not reduce all osteoporosis-related fractures aligns with the ITT findings from SCOOP, ROSE, and SOS. • Their statement that screening may reduce hip fractures reflects the SCOOP trial result (HR 0.72, 84% CI 0.59–0.89; p=0.002), the only statistically significant hip fracture effect among the 3 trials. • The summary of ROSE per-protocol findings is consistent with the trial's reported PP-1 and PP-2 analyses. Conclusion consistent with the data discussed
Gates 2023 (57)	Screening in primary care with clinical FRAX, followed by BMD assessment for increased risk, among selected females aged 65 years and older probably leads to a small reduction in the risk of clinical fragility fracture and hip fracture compared to no screening. Limited direct evidence on harms of screening was available.	The review's conclusions are consistent with the meta-analytic estimates it presents. • For clinical fragility (major osteoporotic) fractures, the pooled estimate (HR 0.93, 95% CI 0.87–0.99) [5.9 or 11.8 fewer per 1000] indicates a small but statistically significant

	SCREENING PROBABLY EFFECTIVE	reduction, supporting the authors' interpretation. - For hip fractures, the meta-analysis (HR 0.80, 95% CI 0.71–0.91) [6.2 or 4 fewer per 1000] also shows a statistically significant reduction, supporting the authors' interpretation. - The statement that there was limited direct evidence on harms is accurate. Conclusions consistent with data presented
Merlijn 2020 (59)	This meta-analysis showed that population screening is effective to reduce osteoporotic fractures and hip fractures. Implementation of screening in older women should be considered as serious option to prevent osteoporotic fractures, especially hip fractures SCREENING EFFECTIVE	The conclusions regarding osteoporotic fractures are not supported by the meta-analytic estimate, whereas the conclusions regarding hip fractures are supported. - For osteoporotic fractures, the pooled estimate (HR 0.95, 95% CI 0.89–1.00) does not demonstrate a statistically significant reduction. The statement that screening programmes reduce all osteoporotic fractures is therefore not consistent with the data presented. - For hip fractures, the pooled estimate (HR 0.80, 95% CI 0.71–0.91) indicates a statistically significant reduction, making this conclusion consistent with the evidence. Conclusions consistent with data for hip fractures but not for osteoporotic fractures But:

		For major osteoporotic fractures the pooled estimate (HR 0.91, 95% CI 0.84–0.98) indicates a statistically significant reduction, making this conclusion consistent with the evidence. Thus, the apparent inconsistency arises from the use of the term osteoporotic fractures in the review conclusions, which extends beyond the statistically significant findings for major osteoporotic fractures
Kahwati 2025 (71)	Screening in higher-risk women 65 years or older was associated with a small absolute risk reduction in hip and major fractures compared with usual care. SCREENING PROBABLY EFFECTIVE	The review's conclusions are consistent with the meta-analytic estimates it presents. - For major osteoporotic fractures, the pooled estimate (HR 0.94, 95% CI 0.88–0.99) [6 fewer per 1000] indicates a small but statistically significant reduction in clinical fragility fracture. - For hip fractures, the pooled estimate (HR 0.83, 95% CI 0.73–0.93) [6 fewer per 1000] supports a small statistically significant reduction in hip fracture risk Conclusions consistent with data presented
EUnetHTA OTCA19 Assessment Team 2019 (56)	The main finding of this Rapid REA is that screening for osteoporosis in postmenopausal women offers no (or only a small) advantage with regard to symptomatic fractures. Since the available moderate quality studies do not show screening to have any effect on the incidence of symptomatic fractures, there is probably	The reviews conclusions are consistent with the trial level and meta-analysis estimates it presents. At the trial level, only one statistically significant numerical trial level estimate was presented for hip fractures from the SCOOP trial, HR 0.72 (95% CI:0.59, 0.89). Notably

	little or no benefit to screening for osteoporosis in postmenopausal women SCREENING NOT EFFECTIVE	ITT data for ROSE were considered for the ROSE trial (not PP-1). The meta-analysis result for osteoporotic fractures was not statistically significant, HR 0.99 (95% CI: 0.94-1.04). Conclusions consistent with data presented
--	---	--

Abbreviations: BMD, bone mineral density; CI, confidence interval; FRAX, Fracture Risk Assessment Tool; HR, hazard ratio; ITT, intention-to-treat; REA, Rapid Evidence Assessment; RR, relative risk; PP, per-protocol.

Review summary

This assessment included 6 reviews relevant to osteoporosis screening, spanning 3 systematic reviews with meta-analysis, a rapid relative effectiveness assessment, a synthesis report and a structured scoping review. Across reviews, objectives and PICOS frameworks were broadly aligned. Most evaluated screening pathways combining clinical risk assessment (often FRAX) with DXA where indicated, followed by treatment for those at elevated risk, compared with usual care or no screening. Despite this broad alignment, author conclusions were mixed, ranging from no clear benefit to probable or effective reductions in fractures in selected populations.

The reviews largely drew on the same underlying evidence base centred on 3 pivotal European RCTs (ROSE (49), SCOOP (50) and SOS (53)). Differences in included studies in addition to the 3 RCTs were mainly explained by search dates, comparator requirements and how the screening pathway was defined (for example, whether BMD measurement by DXA was required, and whether additional vertebral imaging was permitted). These scope and inclusion differences explain some variation in narrative emphasis, but they do not appear to account for the overall pattern of conclusions, as pooled estimates and direction of effects were generally consistent across the meta-analysing reviews.

A potential source of variation appears to be how reviews handled numerical data and in particular, how they interpreted clinical importance. One review (Gates 2023) used hazard ratios derived from time-to-event models, whereas 3 others relied on dichotomous data (57). Time-to-event estimates are methodologically preferable for fracture outcomes, as they account for censoring and differing follow-up time and better reflect the gradual accumulation of benefit observed in the Kaplan–Meier curves. Nonetheless, despite differing analytical inputs, pooled effects were similar across reviews, suggesting that the analytical approach is unlikely to explain the differences in conclusions.

Rather, the divergence in author conclusions appears to be driven by differences in emphasis on relative effects and the baseline risks used to translate relative effects into absolute risk differences. Reviews that foregrounded absolute reductions tended to frame benefits as small, particularly when population-relevant baseline risks were used, whereas others placed greater weight on statistically significant relative effects and interpreted screening as effective. This distinction is important for policy, because identical relative effects can translate into materially different absolute benefits depending on baseline fracture risk.

Finally, the structured scoping review (Auais 2023) provides a plausible explanation for why conclusions may appear less favourable when meta-analysis is not undertaken (55). Although its narrative summary was consistent with the individual trial reports, it could not capture the way modest, non-significant trial-level effects may become statistically significant when pooled. This highlights how the choice of synthesis method and the extent of quantitative pooling can influence the framing of the evidence, even when the underlying studies are similar.

Independent assessment of the evidence base

While all reviews addressed the effectiveness of osteoporosis screening, some were more informative for policy decision-making than others. In particular, the more recent reviews by Gates 2023 (57) and Kahwati 2025 (71) incorporated the latest trial

evidence and presented absolute effect estimates for fracture outcomes in defined populations, facilitating assessment of clinical importance. Notably the populations in the 2 pivotal trials included in the reviews (ROSE and SCOOP) were women aged ≥ 65 years and therefore relevant for whole population screening (research question 1a). SOS included women (aged 65-90 years) with ≥ 1 clinical risk factor⁴ for fractures and therefore is the only study that was directly relevant for targeted screening (research question 1b). However, SOS did not report analyses stratified by individual risk factors, limiting its ability to inform which targeted screening criteria might be most effective in practice. In addition, ROSE included a sub-group analysis of high-risk women which may represent a targeted screening strategy for women with FRAX $\geq 15\%$. Despite the differences in populations, the reviews pooled data across all 3 trials. Furthermore, differences in recruitment, baseline fracture risk, and the timing of risk assessment relative to randomisation mean they do not estimate an equivalent screen-versus-no-screen comparison, limiting the interpretability of combined effect estimates.

Our independent assessment of the meta-analytic evidence across the reviews (Table 4) indicates that osteoporosis screening is not associated with a reduction in all fractures or all-cause mortality. However, screening linked to subsequent treatment is associated with statistically significant reductions in major osteoporotic fractures (including hip, vertebral, wrist, and humerus fractures) and in hip fractures specifically. As the pooled estimates are driven largely by the 2 population-based trials (ROSE and SCOOP), with SOS contributing data from a higher-risk subgroup, they are most applicable to whole-population screening in women aged ≥ 65 years. At the trial level, effects were modest. ROSE, which provides the clearest offer-of-screening comparison in an unselected population (women undergo baseline FRAX after randomisation), did not demonstrate statistically significant fracture reductions on an ITT basis, and most meta-analyses relied on the ROSE PP-1 population (women with sufficient information to calculate baseline FRAX), in which analyses suggested somewhat more favourable effects for screening. SCOOP reported the only statistically significant reduction in an outcome- which was hip fractures (HR, 0.72 [95% CI, 0.59-0.89]) (50).

SOS did not report statistically significant reductions in fracture outcomes; based on this study our independent assessment suggests insufficient evidence to support the effectiveness of targeted osteoporosis screening. The effects might be expected to be larger in a risk-enriched population as indicated by the ROSE PP-2 subgroup analysis; the null findings from SOS may partly reflect suboptimal participation and adherence, as well as the waiting-list design whereby control participants with indications for DXA, vertebral fracture assessment (VFA) and blood tests, were informed and advised to seek assessment through usual care, limiting separation between groups. On the other hand, effects may have been overestimated as all women in the screening arm received risk assessment as well as VFA and BMD.

Overall, pooled findings reflect the accumulation of small effects across trials and should be interpreted cautiously. Further, the underlying evidence base remains small, comprising 3 pivotal pragmatic RCTs, which limits the certainty and generalisability of pooled estimates. In addition, heterogeneity in trial design, recruitment strategies, and screening pathways, together with contamination of usual

⁴ a previous fracture after age 50 years, a parental hip fracture, low body weight, rheumatoid arthritis, early menopause, malabsorption syndrome, chronic liver disease, type I diabetes mellitus, or immobility

care arms, reduces the extent to which trial results can be directly applied to specific population or targeted screening programme models.

Therefore, based on the findings of the evidence map and umbrella review, no further evidence synthesis work on population or targeted screening for osteoporosis should be commissioned at the present time. The UK NSC will return to screening for osteoporosis in women in 3 years' time.

The UK NSC is the only organisation that makes recommendations on screening programmes in the four UK countries. It is important to distinguish that NICE's guideline recommendation on osteoporosis primarily support decision-making for individual patient care. The nature and strength of evidence underpinning clinical guidance differs from evidence required to justify the implementation of a nationally delivered screening programme.

Future considerations

This review was commissioned to explore whether mixed conclusions across reviews reflected genuine disagreement in the evidence base. Our findings suggest that this is not the case. Although reviews differed in scope, screening definitions and analytical choices, they relied on the same 3 pivotal RCTs and reported broadly aligned trial-level and meta-analytic results. Differences in conclusions appear to arise primarily from how these results were interpreted, rather than from differences in the underlying evidence.

In light of this, key considerations for the UK NSC are:

- **Representativeness of the trial evidence.** Whether the existing trials are sufficiently representative of a true population or targeted screening programme. Contamination of control groups and the use of selective recruitment or risk-enriched samples in the pivotal RCTs limits the applicability of findings to routine screening settings both for population and targeted screening. Meta-analyses have pooled trials across heterogeneous recruitment strategies (unselected population-based versus risk-enriched targeted approaches), which complicates interpretation for specific programme models. Several reviews relied on PP-1 rather than strict ITT estimates from ROSE, which may overstate effects relative to the policy-relevant impact of offering screening. Although SOS was the only trial relevant to targeted screening, the screening pathway included additional diagnostic components (e.g. vertebral fracture assessment and biochemical testing to exclude secondary causes of osteoporosis) and all women in the screening arm received follow on testing, which may limit direct applicability to how targeted screening would be delivered in routine implementation.
- **Clinical relevance of observed effects.** Whether the observed benefits are clinically meaningful as well as statistically significant. Although meta-analyses suggest reductions in selected fracture outcomes (hip and major osteoporotic fractures), these effects were modest, were not consistently observed across all fracture outcomes, and were not accompanied by reductions in all-cause mortality. There is some indication that screening may be effective for some women (SCOOP for hip fractures and ROSE PP-2 for women with FRAX >15%). However, the trial designs provide limited evidence to inform the key elements of a potential screening programme, including who should be offered

screening and which primary outcomes should be used to assess its effectiveness.

- **Need for new trials.** Further synthesis of the same evidence is unlikely to change conclusions. Future progress would depend on new trials explicitly designed to evaluate both (i) population screening in truly unselected women aged ≥ 65 years, and (ii) targeted screening approaches in clearly defined high-risk groups, with targeted screening representing the more immediate research priority, using interventions that are directly applicable to routine implementation with careful separation of screening procedures between arms in both settings to minimise contamination (e.g. restricting FRAX assessment to the screening pathway only).
- **Definition of the screening intervention.** What constitutes the screening intervention in practice? Across trials, there is variation in how the screening pathway is implemented, including differences in the use of risk prediction tools, thresholds for treatment, and the extent of follow-on diagnostic testing. This makes it difficult to identify which specific screening approach is being evaluated and limits the ability to translate trial findings into a clearly defined, implementable programme in routine care.
- **Priority outcomes for decision-making.** Which outcomes should be prioritised for screening? There is inconsistency across trials and reviews in the fracture outcomes used to assess effectiveness (e.g. hip fracture, major osteoporotic fracture, fragility fracture), and no clear consensus on which of these should be considered the most policy-relevant endpoint. This creates uncertainty in interpreting the clinical importance of observed effects and in determining whether screening delivers meaningful benefit.

Pragmatic checklist for assessing review comparability when conclusions differ

This review suggests that mixed conclusions across reviews do not necessarily reflect disagreement in the underlying evidence, but differences in interpretation and framing. In such contexts, a structured rapid assessment of review comparability may provide a more proportionate alternative to a full umbrella review. We summarise the key domains used in this assessment as a pragmatic checklist in Appendix 4.

Appendix 1 — Search strategy for the evidence map

Databases and platforms searched

Systematic literature searches were undertaken using the search developed for the 2019 review (48). The searches were conducted in Medline ALL, Embase, The Cochrane Library, CINAHL, Web of Science, Epistemonikos, CEA Registry, International HTA Database, Dissertations and Theses Global and Google Scholar. The search results were limited to those published from November 2024 to date. The retrieved papers were imported into Endnote and deduplicated. The remaining records were uploaded to Covidence for title and abstract screening (70). The records were divided into 2 sets and each set was screened independently by one of 2 reviewers. Prior to this, a pilot screening was conducted in which both reviewers screened the same 100 records to ensure consistency in application of the eligibility criteria.

Search dates

The searches were run on 28 October 2025. The searches were restricted from November 2024 onwards.

Search strategies

Ovid MEDLINE(R) ALL <1946 to November

1. exp Osteoporosis, Postmenopausal/ or exp *Osteoporosis/ – 53,984
2. *Bone Density/ – 30,328
3. Osteoporotic Fractures/ – 9,666
4. *Fractures, Bone/ – 58,071
5. (osteoporo* or osteopenia).tw. – 107,612
6. ((densit* or loss or mass or strength*) adj2 bone*).tw. – 119,367
7. fractur*.ti. – 187,445
8. 1 or 2 or 3 or 4 or 5 or 6 or 7 – 368,575
9. *Mass Screening/ – 60,580
10. (screen or screening or screened or prescreen*).ti. – 241,363
11. 9 or 10 – 259,039
12. 8 and 11 – 1,973
13. (screen or screening or screened or prescreen*).tw. – 1,125,274
14. (fractur* or bone* or osteop*).tw. – 1,201,178

15. ((screen or screening or screened or prescreen*) adj2 (fractur* or bone* or osteop*)).tw. – 2,550
16. 12 or 15 – 3,581
17. ((dexa or dxa) adj2 scan*).tw. – 3,536
18. dual energy x ray absorptiomet*.tw. – 26,835
19. (fracture* adj2 risk* adj2 assessment*).tw. – 1,908
20. FRAX.tw. – 1,998
21. 17 or 18 or 19 or 20 – 30,585
22. osteop*.tw. – 148,144
23. (((((dexa or dxa) adj2 scan*) or dual energy x ray absorptiomet* or (fracture* adj2 risk* adj2 assessment*) or FRAX) adj2 osteop*)).tw. – 471
24. 16 or 23 – 3,982
25. *Absorptiometry, Photon/ – 4,949
26. osteopor*.ti. – 42,673
27. 25 and 26 – 563
28. 24 or 27 – 4,402
29. (scan* and osteop*).ti. – 164
30. 28 or 29 – 4,511
31. (202411* or 2025*).dt,ez,dp,da. – 1,718,771
32. 30 and 31 – 396

Embase

1. exp postmenopause osteoporosis/ – 17,394
2. *osteoporosis/ – 69,504
3. *bone density/ – 34,677
4. *fragility fracture/ – 11,038
5. *acetabulum fracture/ or fracture/ – 139,971
6. (osteoporo* or osteopenia).tw. – 170,788
7. ((densit* or loss or mass or strength*) adj2 bone*).tw. – 169,955
8. fractur*.ti. – 230,536
9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 – 539,213
10. *mass screening/ – 27,443
11. (screen or screening or screened or prescreen*).ti. – 330,849
12. 10 or 11 – 340,133
13. 9 and 12 – 3,001
14. (screen or screening or screened or prescreen*).tw. – 1,642,109
15. (fractur* or bone* or osteop*).tw. – 1,709,444

16. ((screen or screening or screened or prescreen*) adj2 (fractur* or bone* or osteop*)).tw. – 4,272
17. 13 or 16 – 5,807
18. ((dexa or dxa) adj2 scan*).tw. – 9,911
19. dual energy x ray absorptiomet*.tw. – 36,798
20. (fracture* adj2 risk* adj2 assessment*).tw. – 3,367
21. FRAX.tw. – 5,111
22. 18 or 19 or 20 or 21 – 48,053
23. osteop*.tw. – 229,158
24. (((((dexa or dxa) adj2 scan*) or dual energy x ray absorptiomet* or (fracture* adj2 risk* adj2 assessment*) or FRAX) adj2 osteop*)).tw. – 1,037
25. 17 or 24 – 6,701
26. photon absorptiometry/ – 5,051
27. osteopor*.ti. – 63,840
28. 26 and 27 – 740
29. 25 or 28 – 7,328
30. (scan* and osteop*).ti. – 276
31. 29 or 30 – 7,526
32. limit 31 to conference abstract – 2,398
33. 31 not 32 – 5,128
34. limit 33 to "clinical trial (clinicaltrials.gov)" – 148
35. 33 not 34 – 4,980
36. (202411* or 2025*).dc,dd. – 2,617,070
37. 35 and 36 – 415

CINAHL (EBSCO)

1. (MH "Health Screening" OR TI screen*) – 116,591
2. ((MH "Osteoporosis" OR MH "Bone Density" OR MH "Fractures+" OR (osteoporos* or "bone density" or "bone loss" or osteopen* or fracture*)) OR ((FRAX OR DEXA OR DXA) OR MM "Absorptiometry, Photon")) – 150,218
3. 1 AND 2 Limited: 01/11/24–28/10/25 – 111

The Cochrane Library (Wiley)

1. MeSH descriptor: [Mass Screening] explode all trees – 6,127
2. (screening):ti – 12,568
3. (screening):ab – 79,829
4. 1 or 2 or 3 – 83,758

5. (Osteopor* OR "bone density" OR fracture* OR FRAX OR DEXA OR DXA):ti – 21,032
6. MeSH descriptor: [Osteoporosis] explode all trees – 5,544
7. MeSH descriptor: [Bone Density] explode all trees – 6,033
8. MeSH descriptor: [Fractures, Bone] explode all trees – 9,824
9. 5 or 6 or 7 or 8 – 28,098
10. 4 and 9 – 624

The Web of Science (Clarivate)

1. osteopor* or osteopenia (Title) – 54,308
2. screen or screening or screened or prescreen* (Title) – 332,313
3. 2 AND 1 – 1,452
4. (osteop* NEAR/2 screen) or (osteop* NEAR/2 screened) or (osteop* NEAR/2 screening) or (osteop* NEAR/2 prescreen*) (Abstract) – 1,366
5. 4 OR 3 – 2,458
6. (densit* NEAR/2 bone*) or (loss NEAR/2 bone*) or (mass NEAR/2 bone*) or (strength* NEAR/2 bone*) (Title) – 51,281
7. 6 AND 2 – 372
8. 7 OR 5 – 2,675
9. fractur* (Title) – 232,918
10. 9 AND 2 – 571
11. 10 OR 8 – 3,017
12. Limited to 01/11/24–28/10/25 – 253

Epistemonikos

1. (title:(osteoporosis) OR abstract:(osteoporosis)) AND title:(screen*) — 2024–2025, systematic reviews only – 14
2. title:(screen* AND osteoporos*) — 2025, systematic reviews only – 7
3. title:(screen* AND osteoporos*) — 2024, systematic reviews only – 5
4. abstract:(screen* AND osteoporos*) — 2025, systematic reviews only – 66
5. abstract:(screen* AND osteoporos*) — 2024, systematic reviews only – 60

CEA Registry

1. Basic search: Osteoporosis — limited to 2024–2025 – 2

Google Scholar via Publish or Perish, 2024-25 (Limited to first 200 papers. Citations and patents excluded)

1. Osteoporosis and Screening – 200

2. Osteoporosis and Predict and Scan – 200

Numbers of results for each database and question

Table 1. 1: Numbers of records by database searched

Database	Total no. of references	No. of references post-duplication
Medline ALL	396	396
Embase, 1947-	415	94
CINAHL	111	45
Cochrane Reviews	0	0
Cochrane Central	51	44
Web of Science	253	47
Epistemonikos	152	106
CEA Registry	2	2
Google Scholar via Publish or Perish	400	216
TOTAL	1780	950

Inclusions and exclusions

Studies that satisfied the following criteria listed in the tables were included:

Table 1. 2 Question 1a. What is the volume and type of evidence exploring the clinical benefit of whole population screening in women over 65, in reducing osteoporotic fractures compared to standard care?

	Inclusion criteria	Exclusion criteria
Population	Women > 65 years	Male or mixed-gender samples (unless female sample data can be separated) Samples of women under 65 Samples where either risk factors or a mean age of 65 at study entry is not explicit
Target condition	Osteoporosis	Not applicable
Intervention/exposure	Any population or targeted screening ^a intervention that involves any combination of risk	Not applicable

	Inclusion criteria	Exclusion criteria
	assessment and DXA designed to identify women that have, or are at risk of, developing osteoporosis.	
Comparator	Standard care/ no screening	Not applicable
Outcomes	Prevention of osteoporosis: • Osteoporotic fracture: • Vertebral • Non vertebral • Hip • All cause fracture • Mortality Possible harms of screening / treatment, e.g. • Overdiagnosis/overtreatment • False positive results • Anxiety • Thromboembolic event • Osteonecrosis of the jaw • Gastrointestinal problems	Not applicable
Study designs	Systematic reviews Rapid reviews RCTs Non-RCTs Quasi-RCTs	Letters, reviews, editorials, communications, commentaries, conference abstracts and other grey literature, studies that have been retracted Studies where it cannot be established if the inclusion criteria are met
Geographic focus	No restriction	Not applicable

Abbreviations: DXA, Dual-energy X-ray absorptiometry; RCT, randomised controlled trial

^aUK NSC definitions of targeted and population screening were used:

- Population screening: A nationally delivered (in England, Scotland, Wales and Northern Ireland), proactive (individuals are actively invited for screening) screening programme which aims to improve health outcomes in people with the condition being screened for and/or offer information to enable informed choices. It is offered to a group of people identified from the whole population and defined demographically such as by age or sex.
- Targeted screening: A nationally delivered, proactive screening programme which aims to improve health outcomes in people with the condition being screened for, among groups of people identified as being at elevated/above average risk of a specific condition. Compared to the general population, the people targeted may have higher risk because of lifestyle factors, genetic variants or having another health condition. Targeted screening differs from population screening as it aims to identify groups of people with a higher risk of a specific condition beyond demographics such as age or sex. For example, individuals who smoke are at a higher risk of developing lung cancer regardless of their age or sex.

Table 1. 3: Question 1b. What is the volume and type of evidence exploring the clinical benefit of targeted screening of women aged under 65 with risk factors for osteoporosis, in reducing osteoporotic fractures compared to standard care?

	Inclusion criteria	Exclusion criteria
Population	Women < 65 years with risk factors for osteoporosis: • Previous fragility fracture • Current use or frequent use of oral or systemic glucocorticoids • History of falls • Family history of hip fracture • Other causes of secondary osteoporosis • Smoking • Alcohol intake of more than 14 units per week • Menopause	Male or mixed-gender samples (unless female sample data can be separated) Samples of women under 65 where risk factors are not explicit Samples where either risk factors or a mean age of 65 at study entry is not explicit
Target condition	Osteoporosis	Not applicable
Intervention/exposure	Any population or targeted screening ^a intervention that involves any combination of risk assessment and DXA designed to identify women that have, or are at risk of, developing osteoporosis.	Not applicable
Comparator	Standard care/ no screening	Not applicable
Outcomes	Prevention of osteoporosis: • Osteoporotic fracture: • Vertebral	Not applicable

	Inclusion criteria	Exclusion criteria
	• Non vertebral • Hip • All cause fracture • Mortality Possible harms of screening / treatment, e.g. • Overdiagnosis/overtreatment • False positive results • Anxiety • Thromboembolic event • Osteonecrosis of the jaw • Gastrointestinal problems	
Study designs	Systematic reviews Rapid reviews RCTs Non-RCTs Quasi-RCTs	Letters, reviews, editorials, communications, commentaries, conference abstracts and other grey literature, studies that have been retracted Studies where it cannot be established if the inclusion criteria are met
Geographic focus	No restriction	Not applicable

Abbreviations: DXA, Dual-energy X-ray absorptiometry; RCT, randomised controlled trial

^aUK NSC definitions of targeted and population screening were used:

- Population screening: A nationally delivered (in England, Scotland, Wales and Northern Ireland), proactive (individuals are actively invited for screening) screening programme which aims to improve health outcomes in people with the condition being screened for and/or offer information to enable informed choices. It is offered to a group of people identified from the whole population and defined demographically such as by age or sex.
- Targeted screening: A nationally delivered, proactive screening programme which aims to improve health outcomes in people with the condition being screened for, among groups of people identified as being at elevated/above average risk of a specific condition. Compared to the general population, the people targeted may have higher risk because of lifestyle factors, genetic variants or having another health condition. Targeted screening differs from population screening as it aims to identify groups of people with a higher risk of a specific condition beyond demographics such as age or sex. For example, individuals who smoke are at a higher risk of developing lung cancer regardless of their age or sex.

Table 1. 4: Question 2. What is the volume of evidence exploring the cost-effectiveness of screening (whole population and targeted) for osteoporosis?

	Inclusion criteria	Exclusion criteria
Population	Adults	Children
Target condition	Osteoporosis	Not applicable
Intervention/exposure	Any population or targeted screening ^a intervention that involves any combination of risk assessment and DXA designed to identify women that have, or are at risk of, developing osteoporosis.	Not applicable
Comparator	Any comparator	Not applicable
Outcomes	Incremental cost-effectiveness ratio (ICER) Quality-adjusted life years (QALYs), Cost per life-year gained Cost per quality-adjusted life years (DALYs) averted Net health or monetary benefit Overall costs from a healthcare and/or societal perspective, Health-related quality of life (e.g. EQ-5D, SF-36)	Not applicable
Study designs	Cost-effectiveness Analysis Cost-utility analysis Cost-minimisation analysis/ cost comparison Cost-consequence analysis Cost-of-Illness Cost-benefit analysis Budget impact	Studies not reporting cost effectiveness in population screening
Geographic focus	No restriction	Not applicable

Abbreviations: DALY, Disability-adjusted life year; DXA, Dual-energy X-ray absorptiometry; EQ-5D, EuroQol 5-Dimensions; ICER, Incremental cost-effectiveness ratio; QALY, Quality-adjusted life year; RCT, Randomised controlled trial; SF-36, Short Form (36) Health Survey.

^aUK NSC definitions of targeted and population screening were used:

- Population screening: A nationally delivered (in England, Scotland, Wales and Northern Ireland), proactive (individuals are actively invited for screening) screening programme which aims to improve health outcomes in people with the condition being screened for and/or offer information to enable informed choices. It is offered to a group of people identified from the whole population and defined demographically such as by age or sex.
- Targeted screening: A nationally delivered, proactive screening programme which aims to improve health outcomes in people with the condition being screened for, among groups of people identified as being at elevated/above average risk of a specific condition. Compared to the general population, the people targeted may have higher risk because of lifestyle factors, genetic variants or having another health condition. Targeted screening differs from population screening as it aims to identify groups of people with a higher risk of a specific condition beyond demographics such as age or sex. For example, individuals who smoke are at a higher risk of developing lung cancer regardless of their age or sex.

Appendix 2 — Abstract reporting

Question 1a

Citation 1

Kahwati 2025 (71)

Study type

This was a systematic evidence review conducted for the US Preventive Services Task Force (USPSTF), including RCTs, systematic reviews, and diagnostic and predictive accuracy studies.

Objectives

To review and synthesise evidence on osteoporosis screening to inform USPSTF recommendations on the prevention of fractures.

Components of the study

The review included evidence from PubMed, Embase, the Cochrane Library, and trial registries until January 9, 2024, with surveillance until July 31, 2024. Eligible studies included RCTs and systematic reviews of screening, pharmacotherapy studies for primary osteoporosis, and studies assessing predictive or diagnostic accuracy. Three RCTs and 3 systematic reviews evaluated screening in older, higher-risk women, primarily using 2-stage screening approaches involving fracture risk assessment tools with follow-up BMD testing.

Outcomes reported

Screening was associated with reduced hip fractures (pooled RR 0.83, 95% CI 0.73-0.93) and major osteoporotic fractures (pooled RR 0.94, 95% CI 0.88-0.99), corresponding to 5-6 fewer fractures per 1,000 individuals screened. Risk assessment tools showed poor to modest discriminative accuracy (area under curve [AUC] generally 0.60-0.80). Pharmacological treatment with bisphosphonates or denosumab was associated with reduced hip fractures compared with placebo, with no statistically significant increase in adverse events.

Conclusions

The authors conclude that screening in higher-risk women aged 65 years or older results in a small absolute reduction in hip and major osteoporotic fractures compared with usual care. Evidence was insufficient for screening using BMD alone, for screening in men or younger women, and for determining optimal screening intervals.

Question 1b

No citations included for Question 1b.

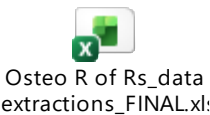
Question 2

No citations included for Question 2.

Appendix 3 — Data extraction and risk of bias assessment

Full data extraction sheet

The data extraction and risk of bias assessment for the reviews, and the data extraction and quality assessment for the primary trials, are provided in the embedded sheet:



Summary of risk of bias

Trials

Figure 2.1. Summary of risk assessment using the Cochrane Risk of Bias 2 tool (RCTs)

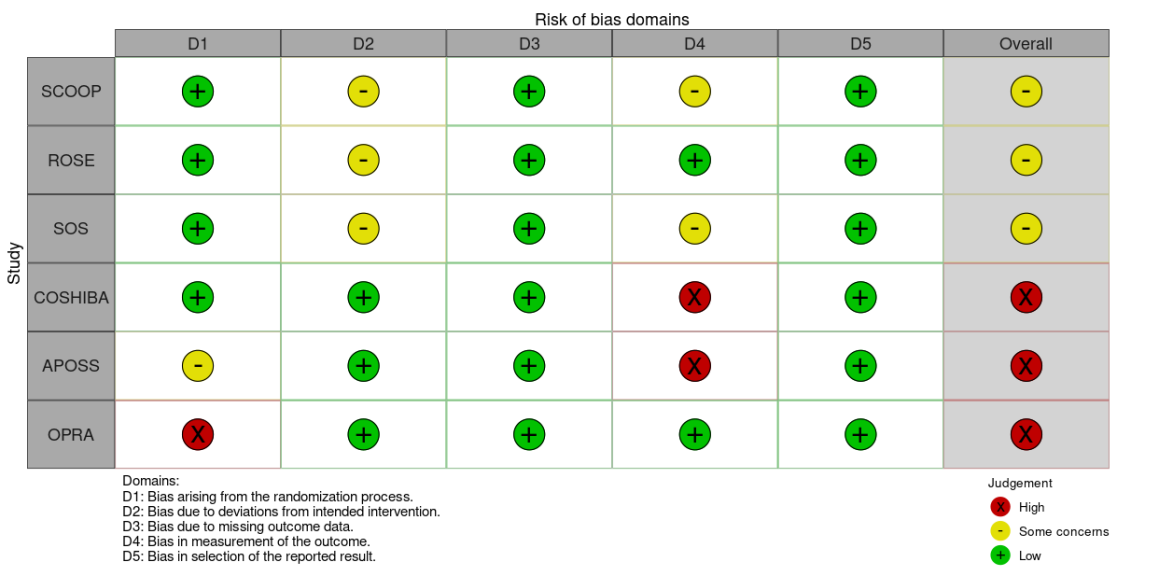
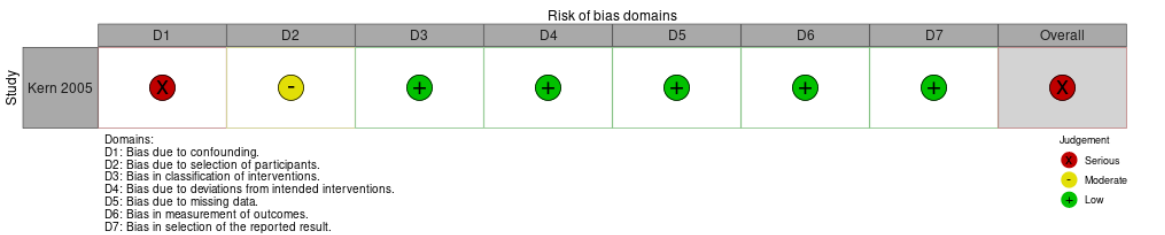


Figure 3.2. Summary of risk assessment using the ROBINS-I tool (nonrandomised studies)



Systematic reviews

Figure 4.3. Summary of risk assessment using ROBIS

Study	Risk of bias				Overall
	D1	D2	D3	D4	
	Merlijn 2019				
	Gates 2023				
	EUnetHTA; 2019				
	Kahwari 2025				
D1: Study eligibility criteria D2: Identification and selection of studies D3: Data collection and study appraisal D4: Synthesis and findings					Judgement High Low

Appendix 4 — Proposed checklist

This checklist summarises the key considerations that emerged from our assessment of why systematic reviews drawing on the same screening trials can reach apparently mixed conclusions. It provides a pragmatic structure for assessing review comparability when differences in conclusions may reflect variation in scope, analytical choices, or interpretation rather than disagreement in the underlying evidence.

It is intended to support proportionate decision making at scoping or pre commissioning stages, helping to determine whether further synthesis is warranted or whether apparent conflicts can be resolved through structured comparison alone. This framework reflects the structured domain-based approach applied in the present umbrella review and can be used to support transparent comparison of reviews across scope, evidence base, analytical choices, and interpretation.

This checklist is provided as a pragmatic structure for assessing review comparability in similar situations.

Step 0. Identify an anchor review

Item	Action
Anchor review selection	Select the most recent and or most policy relevant review, ideally including all known pivotal RCTs, to act as the reference point for comparison.

Step 1. Establish whether reviews are comparable

Question	If Yes	If No
Q1. Does the review draw on substantially the same underlying evidence base as the anchor review?	Evidence base is comparable. Proceed to Q2.	Reviews are not directly comparable due to differences in included trials, scope, or search dates. Conclusions should not be compared directly. No further action required
Q2. Among reviews drawing on the same evidence base, do authors reach consistent overall conclusions?	No substantive conflict. Differences likely reflect emphasis or wording only. Document consistency and no further action is required.	Reviews interpret the same evidence differently. Proceed to conflict analysis (Q3 to Q7).

Step 2. Conflict analysis (reviews differ despite using the same evidence).

Note: Multiple sources of divergence may coexist. Continue through all questions even if differences are identified early.

Question	If Yes	If No
Q3. Do the reviews present comparable numerical data from the trials?	Numerical data are consistent. Continue to Q4.	Record divergence as a data extraction or reporting issue (e.g. different populations, adjusted estimates, errors). Continue to Q4.

Question	If Yes	If No
Q4. Do the reviews agree on which trials are admissible or key (e.g. based on risk of bias assessments)?	Trial inclusion is consistent. Continue to Q5.	Record divergence as a study selection or quality judgement issue. Continue to Q5.
Q5. Are outcomes defined and grouped in similar ways?	Outcomes are comparable. Continue to Q6.	Record divergence as an outcome definition or construction issue. Continue to Q6.
Q6. Do the reviews align in how they synthesise the evidence?	Analytical approach is broadly aligned. Continue to Q7.	Record divergence as an analytical or modelling issue (e.g. effect measures, fixed vs random effects, re analysis). Continue to Q7.
Q7. Do the reviews differ primarily in interpretation of similar numerical findings?	Record divergence as interpretive (e.g. different thresholds for clinical importance, framing of harms). Summarise differences.	No residual interpretive divergence. Apparent differences are likely explained by upstream methodological factors identified in Q3 to Q6.

Summary output

Output	Action
Root cause classification	Map identified divergence(s) to one or more categories: (1) data extraction or reporting differences, (2) study selection or quality judgement differences, (3) outcome definition or construction differences, (4) analytical or modelling choices, (5) interpretive thresholds and framing.
Decision on next steps	Use this structured comparison to determine whether mixed conclusions can be resolved through comparability assessment alone. A full umbrella review is most warranted when reviews draw on the same underlying trials, but divergence persists due to differences in numerical data, trial selection, outcome definitions, or analytical methods (Q3 to Q6). Where divergence is primarily interpretive (Q7), further synthesis is less likely to add value beyond clarifying thresholds for clinical importance. Where an umbrella review is undertaken, the same domains can support transparent comparison across scope, evidence, synthesis, and interpretation.

References

1. National Institute of Arthritis and Musculoskeletal and Skin Diseases. Osteoporosis 2022 [cited 2025 August]. Available from: <https://www.niams.nih.gov/health-topics/osteoporosis>.
2. Yang J, Zeng Y, Yu W. Criteria for osteoporosis diagnosis: a systematic review and meta-analysis of osteoporosis diagnostic studies with DXA and QCT. *EClinicalMedicine*. 2025;83:103244.
3. Ensrud KE, Crandall CJ. Osteoporosis. *Ann Intern Med*. 2024;177(1):ITC1-ITC16.
4. Zhu Z, Yu P, Wu Y, Wu Y, Tan Z, Ling J, et al. Sex Specific Global Burden of Osteoporosis in 204 Countries and Territories, from 1990 to 2030: An Age-Period-Cohort Modeling Study. *The Journal of nutrition, health and aging*. 2023;27(9):767-74.
5. Hernlund E, Svedbom A, Ivergard M, Compston J, Cooper C, Stenmark J, et al. Osteoporosis in the European Union: medical management, epidemiology and economic burden. A report prepared in collaboration with the International Osteoporosis Foundation (IOF) and the European Federation of Pharmaceutical Industry Associations (EFPIA). *Arch Osteoporos*. 2013;8(1):136.
6. Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. *Osteoporos Int*. 2006;17(12):1726-33.
7. Oden A, McCloskey EV, Kanis JA, Harvey NC, Johansson H. Burden of high fracture probability worldwide: secular increases 2010-2040. *Osteoporos Int*. 2015;26(9):2243-8.
8. Boschitsch EP, Durchschlag E, Dimai HP. Age-related prevalence of osteoporosis and fragility fractures: real-world data from an Austrian Menopause and Osteoporosis Clinic. *Climacteric*. 2017;20(2):157-63.
9. Cheng CH, Chen LR, Chen KH. Osteoporosis Due to Hormone Imbalance: An Overview of the Effects of Estrogen Deficiency and Glucocorticoid Overuse on Bone Turnover. *Int J Mol Sci*. 2022;23(3).
10. Pisani P, Renna MD, Conversano F, Casciaro E, Di Paola M, Quarta E, et al. Major osteoporotic fragility fractures: Risk factor updates and societal impact. *World J Orthop*. 2016;7(3):171-81.
11. Johnell O, Kanis JA, Oden A, Sernbo I, Redlund-Johnell I, Petterson C, et al. Mortality after osteoporotic fractures. *Osteoporos Int*. 2004;15(1):38-42.
12. Reginster J-Y, Bruyere O. Osteoporosis: burden, diagnosis, and management. In: Sinclair AJ, Morley JE, Vellas B, Cesari M, Munshi M, editors. *Pathy's Principles and Practice of Geriatric Medicine* 2022. p. 972-8.
13. Moayyeri A, Warden J, Han S, Suh HS, Pinedo-Villanueva R, Harvey NC, et al. Estimating the economic burden of osteoporotic fractures in a multinational study: a real-world data perspective. *Osteoporos Int*. 2023;34(12):2121-32.
14. Clynes MA, Harvey NC, Curtis EM, Fuggle NR, Dennison EM, Cooper C. The epidemiology of osteoporosis. *British Medical Bulletin*. 2020;133(1):105-17.
15. Bouvard B, Annweiler C, Legrand E. Osteoporosis in older adults. *Joint Bone Spine*. 2021;88(3):105135.
16. Harris K, Zagar CA, Lawrence KV. Osteoporosis: Common Questions and Answers. *Am Fam Physician*. 2023;107(3):238-46.
17. US Preventive Services Task Force. Screening for Osteoporosis to Prevent Fractures: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2025;333(6):498-508.
18. National Institute for Health and Care Excellence. Osteoporosis: assessing the risk of fragility fracture, Clinical guideline [CG146] United Kingdom: NICE; updated 2017 [cited 2025 August]. Available from: <https://www.nice.org.uk/guidance/cg146>.

19. Gregson CL, Armstrong DJ, Avgerinou C, Bowden J, Cooper C, Douglas L, et al. The 2024 UK clinical guideline for the prevention and treatment of osteoporosis. *Archives of Osteoporosis*. 2025;20(1):119.
20. Osteoporosis Research Ltd. FRAX Fracture Risk Assessment Tool 2025 [cited 2025 August]. Available from: <https://www.fraxplus.org/calculation-tool/>.
21. Hippisley-Cox J, Coupland C. Derivation and validation of updated QFracture algorithm to predict risk of osteoporotic fracture in primary care in the United Kingdom: prospective open cohort study. *BMJ*. 2012;344:e3427.
22. National Osteoporosis Guideline Group UK (NOGG). Clinical guideline for the prevention and treatment of osteoporosis 2024 [cited 2025 August]. Available from: <https://www.nogg.org.uk/full-guideline>.
23. Todorov G, Brook S, Quah Qin Xian N, Von Widekind S, Freudenthal B, Comninou AN. Comparison of fracture risk calculators in elderly fallers: a hospital-based cross-sectional study. *BMJ Open*. 2022;12(7):e060282.
24. Kanis JA, Johansson H, Harvey NC, McCloskey EV. A brief history of FRAX. *Arch Osteoporos*. 2018;13(1):118.
25. De Laet C, Kanis JA, Oden A, Johansson H, Johnell O, Delmas P, et al. Body mass index as a predictor of fracture risk: a meta-analysis. *Osteoporos Int*. 2005;16(11):1330-8.
26. Kanis JA, Johansson H, Johnell O, Oden A, De Laet C, Eisman JA, et al. Alcohol intake as a risk factor for fracture. *Osteoporos Int*. 2005;16(7):737-42.
27. Kanis JA, Johansson H, Oden A, Johnell O, De Laet C, Eisman JA, et al. A family history of fracture and fracture risk: a meta-analysis. *Bone*. 2004;35(5):1029-37.
28. Kanis JA, Johansson H, Oden A, Johnell O, de Laet C, Melton IL, et al. A meta-analysis of prior corticosteroid use and fracture risk. *J Bone Miner Res*. 2004;19(6):893-9.
29. Kanis JA, Johnell O, De Laet C, Johansson H, Oden A, Delmas P, et al. A meta-analysis of previous fracture and subsequent fracture risk. *Bone*. 2004;35(2):375-82.
30. Kanis JA, Johnell O, Oden A, Johansson H, De Laet C, Eisman JA, et al. Smoking and fracture risk: a meta-analysis. *Osteoporos Int*. 2005;16(2):155-62.
31. Kanis JA, Harvey NC, Johansson H, Oden A, McCloskey EV, Leslie WD. Overview of Fracture Prediction Tools. *J Clin Densitom*. 2017;20(3):444-50.
32. Fordyce G, Hollick R, Black A. What's new in the management of osteoporosis and prevention of fragility fractures? *Orthopaedics and Trauma*. 2024;38(2):99-107.
33. Hippisley-Cox J, Coupland C. Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores. *BMJ*. 2009;339:b4229.
34. Hippisley-Cox J, Coupland C. Validation of QFracture compared with FRAX Analysis prepared for NICE 2011 2011 [cited 2025 Aug]. Available from: <https://www.qresearch.org/media/1033/a-validation-of-qfracture-vs-frax-for-nice-2011-13.pdf>.
35. Genant HK, Engelke K, Fuerst T, Glüer CC, Grampp S, Harris ST, et al. Noninvasive assessment of bone mineral and structure: State of the art. *Journal of Bone and Mineral Research*. 1996;11(6):707-30.
36. Kanis JA. Diagnosis of osteoporosis and assessment of fracture risk. *The Lancet*. 2002;359(9321):1929-36.
37. Krueger D, Tanner SB, Szalat A, Malabanan A, Prout T, Lau A, et al. DXA Reporting Updates: 2023 Official Positions of the International Society for Clinical Densitometry. *Journal of Clinical Densitometry*. 2024;27(1):101437.
38. Shuhart CR, Yeap SS, Anderson PA, Jankowski LG, Lewiecki EM, Morse LR, et al. Executive Summary of the 2019 ISCD Position Development Conference on Monitoring Treatment, DXA Cross-calibration and Least Significant Change, Spinal Cord Injury, Peri-

prosthetic and Orthopedic Bone Health, Transgender Medicine, and Pediatrics. *Journal of Clinical Densitometry*. 2019;22(4):453-71.

39. Rheumatology TT. World Health Organisation (WHO) Classification Criteria for Osteoporosis. Trinidad: Rheumatology TT; 2025 [cited 2026 January]. Available from: <https://rheumatologytt.com/2025/07/30/world-health-organisation-who-classification-criteria-for-osteoporosis/>.
40. NHS. Osteoporosis 2022 [cited 2026 January]. Available from: <https://www.nhs.uk/conditions/osteoporosis/>.
41. Anupama S, Lim SY, Bolster MB. Updates on the Role of DXA in the Evaluation and Monitoring of Osteoporosis. *Current Rheumatology Reports*. 2025;27(1):38.
42. Slart RHJA, Punda M, Ali DS, Bazzocchi A, Bock O, Camacho P, et al. Updated practice guideline for dual-energy X-ray absorptiometry (DXA). *European Journal of Nuclear Medicine and Molecular Imaging*. 2025;52(2):539-63.
43. Palacios S. Medical treatment of osteoporosis. *Climacteric*. 2022;25(1):43-9.
44. Langdahl BL. Overview of treatment approaches to osteoporosis. *British Journal of Pharmacology*. 2021;178(9):1891-906.
45. Ho WC, Chang CC, Wu WT, Lee RP, Yao TK, Peng CH, et al. Effect of Osteoporosis Treatments on Osteoarthritis Progression in Postmenopausal Women: A Review of the Literature. *Curr Rheumatol Rep*. 2024;26(5):188-95.
46. National Institute for Health and Care Excellence. Osteoporosis, Quality standard [QS149] United Kingdom: NICE; 2017 [cited 2025 August]. Available from: <https://www.nice.org.uk/guidance/qs149>.
47. UK National Screening Committee. Adult screening programme: Osteoporosis 2019 [cited 2025 August]. Available from: <https://view-health-screening-recommendations.service.gov.uk/osteoporosis/>.
48. Solutions for Public Health. Screening for osteoporosis. External review against programme appraisal criteria for the UK National Screening Committee 2019 [cited 2025 August]. Available from: <https://view-health-screening-recommendations.service.gov.uk/document/407/download>.
49. Rubin KH, Rothmann MJ, Holmberg T, Hoiberg M, Moller S, Barkmann R, et al. Effectiveness of a two-step population-based osteoporosis screening program using FRAX: the randomized Risk-stratified Osteoporosis Strategy Evaluation (ROSE) study. *Osteoporos Int*. 2018;29(3):567-78.
50. Shepstone L, Lenaghan E, Cooper C, Clarke S, Fong-Soe-Khioe R, Fordham R, et al. Screening in the community to reduce fractures in older women (SCOOP): a randomised controlled trial. *Lancet*. 2018;391(10122):741-7.
51. Peto L, Allaby M. Screening for Osteoporosis in Postmenopausal Women. A report for the UK National Screening Committee 2013 [cited 2025 August]. Available from: <https://share.google/mwV6v9jWldX2wFIPx>.
52. UK National Screening Committee. An evidence map to outline the volume and type of evidence related to screening for osteoporosis for the UK National Screening Committee. 2025.
53. Merlijn T, Swart KMA, van Schoor NM, Heymans MW, van der Zwaard BC, van der Heijden AA, et al. The Effect of a Screening and Treatment Program for the Prevention of Fractures in Older Women: A Randomized Pragmatic Trial. *Journal of Bone and Mineral Research*. 2019;34(11):1993-2000.
54. Kern LM, Powe NR, Levine MA, Fitzpatrick AL, Harris TB, Robbins J, et al. Association between Screening for Osteoporosis and the Incidence of Hip Fracture. *Annals of Internal Medicine*. 2005;142(3):173-81.

55. Auais M, Angermann H, Grubb M, Thomas C, Feng C, Chu CH. The effectiveness and cost-effectiveness of clinical fracture-risk assessment tools in reducing future osteoporotic fractures among older adults: a structured scoping review. *Osteoporos Int.* 2023;34(5):823-40.
56. EUnetHTA Assessment Team. Screening for osteoporosis in the general population. Collaborative Assessment. Diemen (The Netherlands)2019 [cited 2025 August]. Available from: <https://www.eunetHTA.eu>.
57. Gates M, Pillay J, Nuspl M, Wingert A, Vandermeer B, Hartling L. Screening for the primary prevention of fragility fractures among adults aged 40 years and older in primary care: systematic reviews of the effects and acceptability of screening and treatment, and the accuracy of risk prediction tools. *Syst Rev.* 2023;12(1):51.
58. Johnell O, Hertzman P. What evidence is there for the prevention and screening of osteoporosis? Copenhagen, WHO Regional Office for Europe2006 [cited 2025 August]. Available from: <http://www.euro.who.int/document/e88668.pdf>.
59. Merlijn T, Swart KMA, van der Horst HE, Netelenbos JC, Elders PJM. Fracture prevention by screening for high fracture risk: a systematic review and meta-analysis. *Osteoporos Int.* 2020;31(2):251-7.
60. Viswanathan M, Reddy S, Berkman N, Cullen K, Middleton JC, Nicholson WK, et al. Screening to Prevent Osteoporotic Fractures: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *JAMA.* 2018;319(24):2532-51.
61. Condurache CI, Chiu S, Chotiarnwong P, Johansson H, Shepstone L, Lenaghan E, et al. Screening for high hip fracture risk does not impact on falls risk: a post hoc analysis from the SCOOP study. *Osteoporos Int.* 2020;31(3):457-64.
62. McCloskey E, Johansson H, Harvey NC, Shepstone L, Lenaghan E, Fordham R, et al. Management of Patients With High Baseline Hip Fracture Risk by FRAX Reduces Hip Fractures—A Post Hoc Analysis of the SCOOP Study. *Journal of Bone and Mineral Research.* 2018;33(6):1020-6.
63. Parsons CM, Harvey N, Shepstone L, Kanis JA, Lenaghan E, Clarke S, et al. Systematic screening using FRAX((R)) leads to increased use of, and adherence to, anti-osteoporosis medications: an analysis of the UK SCOOP trial. *Osteoporos Int.* 2020;31(1):67-75.
64. Soreskog E, Borgstrom F, Shepstone L, Clarke S, Cooper C, Harvey I, et al. Long-term cost-effectiveness of screening for fracture risk in a UK primary care setting: the SCOOP study. *Osteoporos Int.* 2020;31(8):1499-506.
65. Turner DA, Khioe RFS, Shepstone L, Lenaghan E, Cooper C, Gittoes N, et al. The Cost-Effectiveness of Screening in the Community to Reduce Osteoporotic Fractures in Older Women in the UK: Economic Evaluation of the SCOOP Study. *Journal of Bone and Mineral Research.* 2018;33(5):845-51.
66. Høiberg MP, Rubin KH, Holmberg T, Rothmann MJ, Möller S, Gram J, et al. Use of antiosteoporotic medication in the Danish ROSE population-based screening study. *Osteoporosis International.* 2019;30(6):1223-33.
67. Holmberg T, Möller S, Rothmann MJ, Gram J, Herman AP, Brixen K, et al. Socioeconomic status and risk of osteoporotic fractures and the use of DXA scans: data from the Danish population-based ROSE study. *Osteoporosis International.* 2019;30(2):343-53.
68. Petersen TG, Abrahamsen B, Høiberg M, Rothmann MJ, Holmberg T, Gram J, et al. Ten-year follow-up of fracture risk in a systematic population-based screening program: the risk-stratified osteoporosis strategy evaluation (ROSE) randomised trial. *eClinicalMedicine.* 2024;71.
69. Rothmann MJ, Möller S, Holmberg T, Højberg M, Gram J, Bech M, et al. Non-participation in systematic screening for osteoporosis—the ROSE trial. *Osteoporosis International.* 2017;28(12):3389-99.
70. Veritas Health Innovation. Covidence systematic review software Melbourne [cited 2025 August]. Available from: <http://www.covidence.org>.

71. Kahwati LC, Kistler CE, Booth G, Sathe N, Gordon RDA, Okah E, et al. Screening for Osteoporosis to Prevent Fractures: A Systematic Evidence Review for the US Preventive Services Task Force. *JAMA*. 2025;333(6):509-31.
72. Whiting P, Savovic J, Higgins JP, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016;69:225-34.
73. Sterne JAC, Savovic J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898.
74. Sterne JA, Hernan MA, Reeves BC, Savovic J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919.
75. Clark EM, Gould V, Morrison L, Ades AE, Dieppe P, Tobias JH. Randomized controlled trial of a primary care-based screening program to identify older women with prevalent osteoporotic vertebral fractures: Cohort for skeletal health in Bristol and Avon (COSHIBA). *Journal of Bone and Mineral Research*. 2012;27(3):664-71.
76. 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):561-8.
77. Lacroix AZ, Buist DS, Brenneman SK, Abbott TA, 3rd. Evaluation of three population-based strategies for fracture prevention: results of the osteoporosis population-based risk assessment (OPRA) trial. *Med Care*. 2005;43(3):293-302.