# Screening for Male Osteoporosis: A Systematic Review

April 2022

U.S. Department of Veterans Affairs
Veterans Health Administration
*Health Services Research & Development Service*

**Recommended citation:** Sagalla N, Alexopoulos AS, Gordon AM, et al. Screening for Male Osteoporosis: A Systematic Review. Washington, DC: Evidence Synthesis Program, Health Services Research and Development Service, Office of Research and Development, Department of Veterans Affairs. VA ESP Project #09-010; 2022.

# AUTHORS

Author roles, affiliations, and contributions to the present report (using the CRediT taxonomy) are summarized in the table below.

| Author | Role and Affiliation | Report Contribution |
|---|---|---|
| Nicole Sagalla, MD, MHS | Assistant Professor, Department of Medicine, Division of General Internal Medicine, Duke University School of Medicine<br>Durham, NC | Conceptualization, Investigation, Methodology, Writing – original draft; Writing – review & editing |
| Anastasia-Stefania Alexopoulos, MBBS, MHS | Staff Endocrinologist, Durham VA Health Care System<br>Durham, NC<br><br>Assistant Professor of Medicine, Division of Endocrinology, Duke University Medical Center<br>Durham, NC | Conceptualization, Investigation, Methodology, Writing – original draft; Writing – review & editing |
| Adelaide M. Gordon, MPH | Project Coordinator, Durham Evidence Synthesis Program (ESP) Center<br>Durham, NC<br><br>Research Health Science Specialist, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System<br>Durham, NC | Conceptualization, Data curation, Methodology, Investigation, Project administration, Formal analysis, Writing – original draft, Writing – review & editing |
| Belinda Ear, MPH | Research Assistant, Durham ESP Center<br>Durham, NC<br><br>Research Health Science Specialist, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System<br>Durham, NC | Conceptualization, Data curation, Methodology, Investigation, Project administration, Formal analysis, Visualization, Writing – original draft, Writing – review & editing |
| Vesta C. Nwankwo, MD, MHS | Intern, Department of Internal Medicine, University of Florida<br>Gainesville, FL | Investigation, Validation, Writing – review & editing |
| Harry A Mystakelis MD, MHS | Co-Investigator, Medstar Georgetown University Hospital Department of Pediatrics<br>Washington, DC | Investigation, Validation, Writing – review & editing |
| Dan V. Blalock PhD, MA | Clinical Research Psychologist, Durham Center of Innovation to Accelerate Discovery and | Investigation, Writing – original draft, Writing – review & editing |

| | | |
|---|---|---|
| | Practice Transformation, Durham VA Medical Center<br>Durham, NC<br><br>Assistant Consulting Professor, Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine<br>Durham, NC | |
| Elizabeth Van Voorhees, PhD | Psychologist, Durham VA Medical Center<br>Durham, NC<br><br>Assistant Professor in Psychiatry and Behavioral Sciences, Duke University School of Medicine, Durham, NC | Investigation, Writing – original draft, Writing – review & editing |
| Janet M. Grubber, MSPH | Senior Statistician, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System<br>Durham, NC | Conceptualization, Investigation, Writing – review & editing |
| Richard H. Lee, MD, MPH | Co-Investigator, Durham VA Medical Center, Duke University School of Medicine<br>Durham, NC | Conceptualization, Investigation, Writing – review & editing |
| Matt Crowley MD, MHS | Core Investigator, Durham VA Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System Durham NC<br><br>Associate Professor of Medicine Division of Endocrinology, Department of Medicine, Duke University School of Medicine<br>Durham, NC | Investigation, Methodology, Writing – review & editing |
| Diana Soliman, MD | Fellow, Division of Endocrinology, Diabetes and Metabolism, Duke University School of Medicine<br>Durham, NC | Investigation; Methodology, Writing – review & editing |
| Scott M. Carlson, MD | Fellow, Division of Endocrinology, Diabetes and Metabolism, Duke University School of Medicine<br>Durham, NC | Investigation, Methodology, Writing – review & editing |
| Andrzej Kosinski, PhD | Professor of Biostatistics, Department of Biostatistics and | Formal analysis, Methodology, Writing – review & editing |

| | | |
|---|---|---|
| | Bioinformatics Duke University School of Medicine<br>Durham, NC | |
| Sarah Cantrell, MLIS | Associate Director for Research & Education, Duke University Medical Center Library & Archives, Duke University School of Medicine<br>Durham, NC | Conceptualization, Methodology, Writing – review & editing |
| Karen M. Goldstein, MD, MSPH | Co-director, Durham ESP Center<br>Durham, NC<br><br>Core Investigator, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System<br>Durham, NC<br><br>General Internist, Durham VA Medical Center<br>Durham, NC<br><br>Associate Professor, Department of Medicine, Division of General Internal Medicine, Duke University<br>Durham, NC | Conceptualization, Methodology, Writing – review & editing |
| John W Williams Jr., MD, MHS | Scientific Advisor, Durham ESP Center<br>Durham, NC<br><br>Investigator, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System<br>Durham, NC<br><br>Staff Physician, Durham VA Medical Center<br>Durham, NC<br><br>Professor, Department of Medicine, Division of General Internal Medicine, Duke University<br>Durham, NC | Conceptualization, Methodology, Writing – review & editing |
| Jennifer M. Gierisch PhD, MPH | Co-director, Durham ESP Center | Conceptualization, Investigation, Methodology, Supervision, Formal |

| | | |
|---|---|---|
| | Durham, NC | analysis, Writing – original draft, Writing – review & editing |
| | Core Investigator, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System<br>Durham, NC | |
| | Associate Professor, Department of Population Health Sciences, Duke University School of Medicine<br>Durham, NC | |
| | Associate Professor, Department of Medicine, Division of General Internal Medicine, Duke University School of Medicine<br>Durham, NC | |

This report was prepared by the Evidence Synthesis Program Coordinating Center located at the **Durham VA Medical Center, Durham, NC,** directed by Jennifer M. Gierisch PhD, MPH and Karen M. Goldstein, MD, MSPH and funded by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development.

# PREFACE

The VA Evidence Synthesis Program (ESP) was established in 2007 to provide timely and accurate syntheses of targeted health care topics of importance to clinicians, managers, and policymakers as they work to improve the health and health care of Veterans. These reports help:

- Develop clinical policies informed by evidence;
- Implement effective services to improve patient outcomes and to support VA clinical practice guidelines and performance measures; and
- Set the direction for future research to address gaps in clinical knowledge.

The program comprises three ESP Centers across the US and a Coordinating Center located in Portland, Oregon. Center Directors are VA clinicians and recognized leaders in the field of evidence synthesis with close ties to the AHRQ Evidence-based Practice Center Program. The Coordinating Center was created to manage program operations, ensure methodological consistency and quality of products, interface with stakeholders, and address urgent evidence needs. To ensure responsiveness to the needs of decision-makers, the program is governed by a Steering Committee composed of health system leadership and researchers. The program solicits nominations for review topics several times a year via the program website.

The present report was developed in response to a request from the Office of Specialty Care and Osteoporosis Field Advisory Committee, for the purpose of informing VA practice. The scope was further developed with input from Operational Partners (below), the ESP Coordinating Center, the review team, and the technical expert panel (TEP). The ESP consulted several technical and content experts in designing the research questions and review methodology. In seeking broad expertise and perspectives, divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Ultimately, however, research questions, design, methodologic approaches, and/or conclusions of the review may not necessarily represent the views of individual technical and content experts.

# ACKNOWLEDGMENTS

In designing the study questions and methodology at the outset of this report, the ESP consulted several technical and content experts. Broad expertise and perspectives were sought. Divergent and conflicting opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design, methodologic approaches, and/or conclusions do not necessarily represent the views of individual technical and content experts.

The authors are grateful to Dr. Stacy Lavin and Liz Wing for editorial and citation management support, and the following individuals for their contributions to this project:

### Operational Partners

Operational partners are system-level stakeholders who help ensure relevance of the review topic to the VA, contribute to the development of and approve final project scope and timeframe for completion, provide feedback on the draft report, and provide consultation on strategies for dissemination of the report to the field and relevant groups.

**Leonard Pogach, MD**
*Chief Consultant*
Specialty Care Services

**Robert Adler, MD**
*Chair*
Osteoporosis Field Advisory Committee

**Grant Cannon, MD**
*Chair*
Osteoporosis Field Advisory Committee

## Technical Expert Panel (TEP)

To ensure robust, scientifically relevant work, the TEP guides topic refinement; provides input on key questions and eligibility criteria, advising on substantive issues or possibly overlooked areas of research; assures VA relevance; and provides feedback on work in progress. TEP members are listed below:

**Cathleen Colon-Emeric, MD**
*Associate Director-Clinician, GRECC*
Durham VA Health Care System
Durham, NC

**Karen Hansen, MD, MS**
*Associate professor*
University of Wisconsin
Madison, WI

**Anne Schafer, MD**
*Staff Physician*
San Francisco VA Health Care System
San Francisco, CA

## Peer Reviewers

The Coordinating Center sought input from external peer reviewers to review the draft report and provide feedback on the objectives, scope, methods used, perception of bias, and omitted evidence. Peer reviewers must disclose any relevant financial or non-financial conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The Coordinating Center and the ESP Center work to balance, manage, or mitigate any potential nonfinancial conflicts of interest identified.

# TABLE OF CONTENTS

**Authors** ..... i
**Preface** ..... v
**Acknowledgments** ..... v
**Executive Summary** ..... 1
Introduction ..... 1
Methods ..... 2
Results ..... 3
Discussion ..... 5
**Introduction** ..... 8
**Methods** ..... 10
Topic Development ..... 10
Search Strategy ..... 10
Study Selection ..... 10
Data Abstraction ..... 13
Quality Assessment ..... 14
Data Synthesis ..... 15
Rating the Body of Evidence ..... 15
Peer Review ..... 16
**Results** ..... 17
Literature Flow ..... 17
Evidence Profile ..... 21
Organization of Results ..... 22
Key Question 1: Among males not identified by a history of low-trauma fracture, is there a clinical risk tool (*eg*, FRAX) that identifies patients at highest risk of osteoporosis or major osteoporotic fracture? ..... 23
Quality of Evidence for Key Question 1 ..... 33
Key Question 2: Among male Veterans not identified by a history of low-trauma fracture, is there a tool or combination of risk factors that identify patients at highest risk of osteoporosis or major osteoporotic fracture? ..... 36
Quality of Evidence for Key Question 2 ..... 58
Key Question 3: What system-level interventions improve uptake of osteoporosis screening among people not identified by a history of low-trauma fracture? ..... 62
Quality of Evidence for Key Question 3 ..... 74
**Summary and Discussion** ..... 78
Summary of Evidence by Key Question ..... 78

Prior Systematic Reviews ..... 85
Clinical Policy Implications ..... 87
Limitations ..... 88
Research Gaps/Future Research ..... 90
Conclusions ..... 91
**References** ..... 93

## TABLES

Table 1. Study Eligibility Criteria ..... 11
Table 2. Evidence Profile of Included Studies ..... 21
Table 3. FRAX, Modified FRAX and OST for Assessing Risk of Osteoporosis and/or Fracture in Male Veterans ..... 41
Table 4. Studies Examining Mscore and VA-FARA as One-off Tools to Predict Osteoporosis and/or Fracture in Male Veterans ..... 47
Table 5. Assessment of Individual Risk Factors ..... 52
Table 6. Provider Education ..... 63
Table 7. Provider and Patient Education ..... 65
Table 8. Provider-focused Reminders ..... 67
Table 9. Clinical Decision Support Tool ..... 69
Table 10. Patient Navigation ..... 70
Table 11. Patient Risk Assessment ..... 71
Table 12. Patient Self-referral ..... 72
Table 13. Patient Reminders ..... 74
Table 14. Certainty of Evidence for Main Outcomes of Osteoporosis Risk Assessment Tools ..... 79
Table 15. Certainty of Evidence for Uptake of Osteoporosis Screening by Intervention Type ..... 84
Table 16. Evidence Gaps and Future Research for Studies of Clinical Risk Prediction Tool and Risk Factors Among Male Veterans ..... 90
Table 17. Evidence Gaps and Future Research for Studies of Systems-level Approaches to Improve Uptake of Osteoporosis Screening ..... 90

## FIGURES

Figure 1. Literature Flow Chart: KQ 1 and KQ 2 ..... 18
Figure 2. Literature Flow Chart: KQ 3 ..... 20
Figure 3. FRAX Tool Compared to Major Osteoporotic Fracture ..... 25

Figure 4. FRAX Tool Compared to Hip Fracture ..... 25
Figure 5. FRAX Tool Compared to Osteoporosis ..... 26
Figure 6. OST Tool Compared to Osteoporosis ..... 29
Figure 7. QFracture Tool Compared to Major Osteoporotic Fracture ..... 30
Figure 8. QFracture Tool Compared to Hip Fracture ..... 30
Figure 9. MORES Tool Compared to Osteoporosis ..... 31
Figure 10. Risk of Bias Ratings for the Included KQ 1 Studies ..... 35
Figure 11. Risk of Bias Assessment Across Included KQ 1 Studies ..... 36
Figure 12. Risk of Bias for Included KQ 2 Studies Evaluating Combined Risk Factors for Osteoporosis and/or Fracture in Male Veterans ..... 59
Figure 13. Risk of Bias Across KQ 2 Studies Evaluating Combined Risk Factors for Osteoporosis and/or Fracture in Male Veterans ..... 59
Figure 14. Risk of Bias for Included Studies Evaluating Individual Risk Factors for Osteoporosis and/or Fracture in Male Veterans ..... 60
Figure 15. Risk of Bias for Included Case-Control Study Evaluating Individual Risk Factors for Osteoporosis and/or Fracture in Male Veterans ..... 60
Figure 16. Risk of Bias Across Included KQ 2 Studies Evaluating Individual Risk Factors for Osteoporosis and/or Fracture in Male Veterans ..... 61
Figure 17. Impact of Provider Education on Update of Osteoporosis Screening ..... 64
Figure 18. Impact of Provider Education plus Patient Education on Uptake of Osteoporosis Screening ..... 66
Figure 19. Impact of Provider-focused Reminders on Uptake of Osteoporosis Screening ..... 68
Figure 20. Patient Self-referral on Uptake of Osteoporosis Screening ..... 73
Figure 21. Risk of Bias Ratings for the Included Studies ..... 75
Figure 22. Risk of Bias Assessment Across Included Studies ..... 76
Figure 23. Risk of Bias Ratings for the Interrupted Time-Series Studies ..... 76
Figure 24. Risk of Bias Assessment Across the Interrupted Time-Series Studies ..... 77

**Appendix A. Search Strategies** ..... 100
**Appendix B. KQ 1 and KQ 2 Study Characteristics Table** ..... 121
**Appendix C. KQ 3 Study Characteristics Table** ..... 137
**Appendix D. KQ 1 and KQ 2 Excluded Studies** ..... 145
**Appendix E. KQ 3 Excluded Studies** ..... 179
**Appendix F. Peer Review Comments** ..... 201
**Appendix G. Tools** ..... 211

# EVIDENCE REPORT

## INTRODUCTION

Osteoporosis is a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue, with a consequent increase in bone fragility and susceptibility to fracture. Although studied mainly in postmenopausal women, osteoporosis has been recognized as a prevalent disease in men through similar mechanisms (*ie,* age-related bone loss, hormonal alterations, and other conditions/risk factors associated with bone loss). The lifetime risk of an osteoporotic fracture in men over the age of 50 is between 20% and 30%.[1] Although this is less than the overall prevalence in women, men have higher rates of fracture-related mortality than women (mortality rate of 73.0 in women versus 166.5 in men per thousand person-years).[2]

Primary prevention of osteoporosis is largely sought through screening to identify those at the highest risk of fracture-related morbidity. While screening women for osteoporosis is standard clinical practice, there is uncertainty about the role of screening among men.[3] The Bone Health and Osteoporosis Foundation, the International Society for Clinical Densitometry,[4] and the Endocrine Society[5] recommend screening all men over 70 years of age and younger men with risk factors for osteoporosis. In 2018, however, the United States Preventive Services Task Force (USPSTF) found insufficient evidence to recommend screening men for osteoporosis.[6] Likely as a result of uncertainty in screening recommendations, screening rates for osteoporosis are low among men.[7] In addition to whether or not to screen men for osteoporosis, there is also uncertainty about how to screen men when screening is determined to be warranted. Moreover, most male fragility fractures occur in those with bone mineral density (BMD)-defined osteopenia rather than osteoporosis due to the higher prevalence of osteopenia.[8] In response, fracture risk-assessment tools, such as the FRAX® tool, have been developed to identify those who may not have BMD-defined osteoporosis but are at high risk for fracture. Screening first by fracture risk, rather than by DXA, has been proposed as an alternate means to identify those at increased risk for fracture.[9] It is unknown how effective such fracture risk assessment tools are among men. While men are often identified for osteoporosis screening and treatment because of a low-impact fracture, for primary prevention, it is critical to identify those at risk before clinically relevant effects of osteoporosis emerge.

The issue of screening for osteoporosis among men is particularly pertinent to the Veterans Health Administration (VHA). Veterans of both sexes are at higher risk for osteoporotic fractures,[10] have more chronic medical conditions, have less bone-health knowledge, spend less time exercising, and have more falls than non-Veterans, contributing to a higher risk for bone health problems.[11] It is not known whether screening for osteoporosis and/or increased fracture risk in this population will reduce the future risk of osteoporotic fracture.[10,11] Thus, this report seeks to assess sensitivity/specificity of osteoporosis risk assessment tools among men, individual factors associated with increased risk of osteoporosis among male Veterans, the effectiveness of osteoporosis screening on patient-important outcomes such as screening rates and fracture rates, and system-level approaches for boosting osteoporosis screening among men.

## Key Questions

The key questions (KQs) for this report were:

**KQ 1:** Among males not identified by a history of low-trauma fracture, is there a clinical risk tool (*eg*, FRAX®) that identifies patients at highest risk of osteoporosis or major osteoporotic fracture?

**KQ 2:** Among male Veterans not identified by a history of low-trauma fracture, is there a tool or combination of risk factors that identifies patients at highest risk of osteoporosis or major osteoporotic fracture?

**KQ 3:** What system-level interventions improve uptake of osteoporosis screening among people without a history of low-trauma fracture?

# METHODS

We followed a standard protocol for this review developed in collaboration with operational partners and a technical expert panel. The PROSPERO registration number is CRD42020150830. The protocol was developed prior to the conduct of the review, and there were no significant deviations after registration. Each step was pilot-tested to train and calibrate study investigators. We adhered to the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) guidelines.[12]

## TOPIC DEVELOPMENT

This systematic review was requested by the Office of Specialty Care and the Osteoporosis Field Advisory Committee. The review will be used to identify the current evidence base and its quality to (1) support screening for osteoporosis and increased fracture risk in men, specifically in Veterans, and (2) identify interventions to improve the uptake of screening. Prior systematic reviews on this topic were inadequate for the needs of stakeholders because they did not include recent important studies and did not adequately consider components such as screening by risk factors or fracture risk—which are increased in Veterans—and system-based interventions to improve screening uptake.

## SEARCH STRATEGY

In collaboration with an expert reference librarian, we conducted a primary literature search from inception to June 28, 2019, for KQ 1 and KQ 2, and to July 22, 2019, for KQ 3, of MEDLINE® (via PubMed®), Embase (via Elsevier), and CINAHL (via EBSCO). We subsequently updated the search for KQ 1, KQ 2, and KQ 3 in MEDLINE® on February 23, 2021. We used a combination of database-specific subject headings and selected free-text terms (*eg*, osteoporosis, fracture assessment) to search titles and abstracts (Appendix A). We also conducted hand-searches of the references from select high-quality systematic reviews and exemplar studies identified during the topic development process and as identified by our stakeholders.

Our search strategy was informed by the Cochrane Effective Practice and Organization of Care (EPOC) Group.[13] EPOC criteria were developed to capture both randomized and nonrandomized study designs. All citations were imported into 2 electronic databases (for referencing, EndNote®, Clarivate Analytics, Philadelphia, PA; for data abstraction, DistillerSR; Evidence Partners Inc., Manotick, ON, Canada).

## STUDY SELECTION

Key eligibility criteria for study inclusion in KQ 1 and KQ 2 meta-analyses were (1) study design: cohort, case-control, or cross-sectional; (2) study purpose: evaluation of clinical risk assessment tools (*eg,* FRAX®, Garvan, QFracture®, or Osteoporosis Self-assessment Tool [OST]); (3) study population: not identified by a prior low-trauma fracture; and (4) outcomes: osteoporosis**,** osteopenia (*ie,* BMD T-score between -1.0 and 2.5) with an additional risk factor, or fracture. KQ 2 additionally included studies that examined the association between osteoporosis or major osteoporotic fracture (MOF) and potential independent risk factors among male Veterans only. Studies could qualify for both KQ 1 and KQ 2, and we have highlighted these across both KQ 1 and KQ 2 Results sections.

Importantly, we did not include studies that intentionally recruited men because of a history of low-trauma fracture (*eg,* patients attending a clinic because of a fracture). While this group of men is clinically important to target for screening and treatment, populations recruited in this way may not be representative of the larger target population of men at risk and artificially inflate the fracture rate in the study population. However, studies with participants that happened to have a history of low-trauma fractures were still included. Eligible studies defined osteoporosis and osteopenia based on BMD T-scores (*ie,* BMD T-score ≤ -2.5 and BMD T-score between -1.0 and - 2.5 respectively). The International Society for Clinical Densitometry guidelines recommends using a standard Caucasian female reference for men of all ethnic groups.[14]. The reference groups used to calculate T-scores varied widely and this did not affect study eligibility. The definition of osteoporosis for this review does NOT include clinical osteoporosis diagnosis by prior fracture.

Eligibility criteria for KQ 3 included randomized trials, nonrandomized trials, controlled before-after studies, and interrupted time-series studies evaluating system-level interventions for increasing screening for osteoporosis among men or women not identified because of prior fracture.

We used the artificial intelligence (AI) technology developed as part of the DistillerSR software, called DistillerAI, to assist with screening abstracts.[15] Using prespecified inclusion/exclusion criteria (Table 1), the titles and abstracts of a subset of articles (approximately n = 200) identified through our primary search were classified independently by 2 senior investigators (JMG, NS) for relevance to the KQs. After resolving disagreements between the investigators via consensus or by obtaining a third reviewer's opinion, this set of included and excluded articles was used to train the Distiller AI program.

We used Distiller AI to screen the remaining titles and abstracts and assigned a prediction score of relevance to the study questions. All citations classified with a prediction score ≤ 0.5 underwent screening by a single investigator. Potentially relevant studies included by the investigator or with an AI prediction score > 0.5 underwent full-text screening. The sensitivity of machine-assisted screening was comparable to a single-reviewer screening (78% sensitivity for machine-assisted screening and single-reviewer screening). The specificity of machine-assisted screening was 95% with a 95% CI (0.92-0.97).[16]

At the full-text screening stage, 2 independent investigators agreed on a final inclusion/exclusion decision. Disagreements on eligibility were resolved by consensus or by obtaining a third reviewer's opinion when consensus was not reached.[16]

### Table 1. Study Eligibility Criteria

| Study Characteristic | Include | Exclude |
|---|---|---|
| Population | KQ 1: Adult males<br><br>KQ 2: Adult male Veterans<br><br>KQ 1, KQ 2: Studies with mixed populations of men and women were included if they conducted a subgroup analysis of men only; for | KQ 1, KQ 2:<br>• Children<br>• Other metabolic bone diseases (*eg*, osteogenesis imperfecta, osteomalacia/ rickets, renal osteodystrophy, primary |

| Study Characteristic | Include | Exclude |
|---|---|---|
| | studies that analyzed both men and women together, a first-order approach was taken at full-text review.<br><br>KQ 3: Health care providers, adult patients, health system administrators and/or staff.<br><br>KQ 3: For studies that recruit populations with and without facture histories, 80% of recruited study population should have no prior identified low-trauma fracture. | hyperparathyroidism, Paget's, osteopetrosis)<br>• History of low-trauma fractures[a]<br><br>KQ 3: Children |
| Intervention | KQ 1: Clinical risk assessment or fracture risk prediction tools such as Fracture Risk Assessment (FRAX®), Garvan Fracture Risk Calculator (FRC), Q-fracture, Osteoporosis Self-assessment Tool (OST).<br><br>KQ 2: Risk factors for osteoporosis (*eg*, medication use, smoking, body mass index) or clinical risk assessment or fracture risk prediction tools.<br><br>KQ 3: System-level approaches targeting provider behaviors or systems operations to optimize uptake of osteoporosis screening:<br>• Clinical and patient reminder systems<br>• Bone health clinics<br>• Provider education<br>• Targeted/tailored or bidirectional patient education such as an interactive voice response (IVR) assessing individual risk scores or system-level algorithm deployed for patient identification<br>• Remote consultation (*eg*, ECHO[17])<br>• Nurse/physician/pharmacist-led interventions<br>• Clinician incentives<br>• Academic detailing<br>• Patient self-referral system | KQ 1, KQ 2: Drug treatment trials; diagnostic testing in symptomatic populations<br><br>KQ 1: Independent risk factors, additional imaging technologies<br><br>KQ 3: Generic patient or health education that has not been customized on individual patient factors such as age, screening history, or risk (*eg,* generic mailed pamphlet, mass awareness campaigns) |
| Comparator | KQ 1, KQ 2: Other risk assessment tools, bone mineral density testing via validated approach (*eg*, dual-energy x-ray absorptiometry [DXA]).<br><br>KQ 3: Usual care, other system-level approaches, patient-focused interventions. | Studies with no comparator |
| Outcomes | KQ 1, KQ 2: Fracture rates; BMD with osteoporosis (T-score ≤ -2.5) or osteopenia (T-score between -1.0 and -2.5) plus additional risk factor. | None |

| Study Characteristic | Include | Exclude |
|---|---|---|
| | KQ 3: Fracture rates, osteoporosis screening rates. | |
| Timing | KQ 1, KQ 2: Any timing<br><br>KQ 3: Longitudinal, prospective | KQ 3: Cross-sectional |
| Setting | KQ 1, KQ 2, KQ 3:<br>• Outpatient general medical settings (*eg*, geriatrics, family medicine, general internal medicine, integrative medicine, urgent care, emergency departments).<br>• Inpatient health care setting. | KQ 3: Non-health care setting (*eg*, churches, pharmacies not integrated into a health care setting, senior centers) |
| Design[b] | KQ 1, KQ 2: Cohort studies, case-control studies, cross-sectional studies.<br><br>KQ 1, KQ 2, KQ 3: EPOC criteria studies[b] that have prospective data collection:<br>• Randomized trials<br>• Nonrandomized trials<br>• Controlled before-after studies<br>• Interrupted time-series studies or repeated measures studies<br><br>KQ 1, KQ 2, KQ 3: Patient-level meta-analysis | • Self-described pilot studies without adequate power to assess impact of intervention on outcomes<br>• Studies of small sample sizes (n < 100)<br>• Not a clinical study (*eg*, editorial, non-systematic review, letter to the editor)<br>• Uncontrolled clinical study<br>• Qualitative studies<br>• Clinical guidelines<br>• Systematic reviews (only to be scanned for relevant primary studies) |
| Countries | OECD[c] | Non-OECD |
| Language | English abstract | No English abstract |
| Years | Any | None |
| Publication types | Full publication in a peer-reviewed journal | Letters, editorials, reviews, dissertations, meeting abstracts, protocols without results |

*Abbreviations* BMD=bone mineral density; DXA=dual-energy x-ray absorptiometry; FRAX=Fracture Risk Assessment tool; FRC=Fracture Risk Calculator; IVR=interactive voice response; KQ=Key Question; MOST=male osteoporosis screening tool; MORES=male osteoporosis risk estimation score; OECD=Organization for Economic Co-operation and Development; OST=Osteoporosis Self-assessment Tool

[a] Special populations with high risk for osteoporosis were considered for exclusion based on screening guidelines from National Osteoporosis Society and Endocrine Society.

[b] See Cochrane EPOC criteria for definitions and details.[13]

[c] Organization for Economic Co-operation and Development includes Australia, Austria, Belgium, Canada, Chile, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States.

## DATA ABSTRACTION

Data from published reports were abstracted into a customized DistillerSR database by 1 reviewer and over-read by a second reviewer. Disagreements were resolved by consensus or by

obtaining a third reviewer's opinion when consensus was not reached. Data elements included descriptors to assess applicability, quality elements, intervention/exposure details, and outcomes.

Key characteristics abstracted included patient descriptors (*eg*, age, race, Veteran status, comorbidities). For KQ1 and 2, risk assessment tool or risk factors, cutpoints reported per tool for sensitivity and specificity analysis *(eg*, OST score < 2), reference population used for BMD T-score calculation *(eg*, race and gender specific NHANES III), and region of interest for DXA scan (*eg,* lumbar spine, femoral neck, nondominant forearm) were abstracted. Abstracted study characteristics for KQ 3 included intervention details (*eg*, intervention target, duration/intensity, key intervention components), comparator, and outcomes, as described in Table 1. Multiple reports from a single study were treated as a single data point, prioritizing results based on the most complete and appropriately analyzed data. When critical data were missing or unclear in published reports, we requested supplemental data from the study authors.

For KQ 1 and KQ 2, outcomes of interest were primarily reported as test characteristics of screening tools and were abstracted as reported, including area under the curve (AUC), sensitivity/specificity, odds ratios, and observed/expected ratios. AUC conveys the degree to which a tool can discriminate between 2 clinical states (at risk or not at risk),[18] where an AUC of 0.5 is considered no better than chance and AUC of 1 perfectly distinguishes between at risk and not at risk. Within these limits, an AUC of 0.7 to 0.8 is acceptable discrimination and 0.8 to 0.9 is excellent.[19] Observed/expected ratios or odds ratios were abstracted for those studies evaluating individual risk factors or that did not report AUCs or sensitivity and specificity. The observed/expected ratio indicates the extent to which a given population experiences the condition in question (*ie*, osteoporosis or fracture) in relation to what would be expected based on the prediction from a given risk assessment tool.

For details of study characteristics, see Appendices B and C. Appendices D and E list excluded studies and the reason for exclusion.

## QUALITY ASSESSMENT

Quality assessment was done by the investigator abstracting or evaluating the included article and was over-read by a second, highly experienced investigator. Disagreements were resolved by consensus between the 2 investigators or, when needed, by arbitration by a third investigator.

We used the Cochrane EPOC risk of bias (ROB) tool for KQ 3, which is applicable to randomized, nonrandomized, controlled before-after, and interrupted time-series studies.[13] These criteria are adequacy of randomization and allocation concealment; comparability of groups at baseline; blinding; completeness of follow-up and differential loss to follow-up; whether incomplete data were addressed appropriately; validity of outcome measures; protection against contamination; selective outcomes reporting; and conflict of interest. Summary risk of bias ratings include "low risk of bias," "unclear risk of bias," and "high risk of bias."

We assigned a summary risk of bias score ("low risk of bias" or "at risk of bias") to individual studies using the QUADAS-2 for Diagnostic Accuracy Studies.[20] The criteria are patient selection, concerns about the index test, the gold standard, and patient flow and timing. For cohort and case-control studies, we used the modified Newcastle-Ottawa scales.[21] This scale includes quality assessment criteria for selection of cases and controls, comparability of cases

and controls, and ascertainment of exposure (or outcome as relevant). Summary risk of bias rating for the Newcastle-Ottawa scales include "low risk of bias," "unclear risk of bias," and "high risk of bias."

## DATA SYNTHESIS

We summarized the primary literature using relevant data abstracted from the eligible studies. Summary tables describe the key study characteristics of the primary studies: study design, patient demographics, and details of the intervention and comparator, risk assessment tool or risk factors. We then determined the feasibility of completing a quantitative synthesis (*ie*, meta-analysis) to estimate summary effects. For meta-analyses, feasibility depends on the volume of relevant literature (*ie,* at least 3 studies reporting the same outcome), conceptual and statistical homogeneity (*ie,* $I^2 < 90\%$) of the studies, and completeness of results reporting.

We aggregated outcomes when there were at least 3 studies with the same outcome, based on the rationale that 1 or 2 studies do not provide adequate evidence for summary effects. For KQ 3, we grouped outcomes into similar intervention types (*eg*, patient-focused, provider-focused). When meta-analyses were feasible, we conducted them stratified by study design (randomized vs nonrandomized). Studies reported dichotomous outcomes and continuous outcomes. Diagnostic text accuracy outcomes were combined using AUC/ROC and sensitivity and specificity where possible. Similarly, dichotomous outcomes were combined using risk ratio or odds ratio. For analyses with few studies ($n < 20$), we used the Knapp Hartung approach[22] to adjust the standard errors of the estimated coefficients. Sensitivity analyses included analyses that omit studies with patients at increased risk for osteoporosis or of interventions (KQ 3) of varying complexity. We evaluated for statistical heterogeneity using visual inspection and Cochran's Q and $I^2$ statistics. When the $I^2$ test indicated considerable heterogeneity (*ie,* $> 90\%$), we did not present summary estimates, based on the rationale that if 90% of the variability is attributed to study differences, the summary estimate cannot be meaningfully interpreted.

When a quantitative synthesis was not feasible, we narratively analyzed the data. We gave more weight to the evidence from higher-quality studies with more precise estimates of effect. The narrative synthesis focused on documenting and identifying patterns in efficacy across risk prediction tools (KQ 1 and KQ 2) and interventions (KQ 3), comparators, and outcome categories. We analyzed potential reasons for inconsistency in treatment effects across studies by evaluating differences in the study population, intervention, comparator, and outcome definitions.

## RATING THE BODY OF EVIDENCE

The certainty of evidence (COE) for each KQ 1 and KQ 3 was assessed using the partially contextualized approach described by the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) working group.[23] In brief, this approach requires the assessment of 4 domains: risk of bias, consistency, directness, and precision. Additional domains to be used when appropriate are coherence, dose-response association, impact of plausible residual confounders, strength of association (magnitude of effect), and publication bias. These domains were considered qualitatively, and a summary rating was assigned after discussion by 2 investigators (JG, AG) as high, moderate, low, or very low COE.

## PEER REVIEW

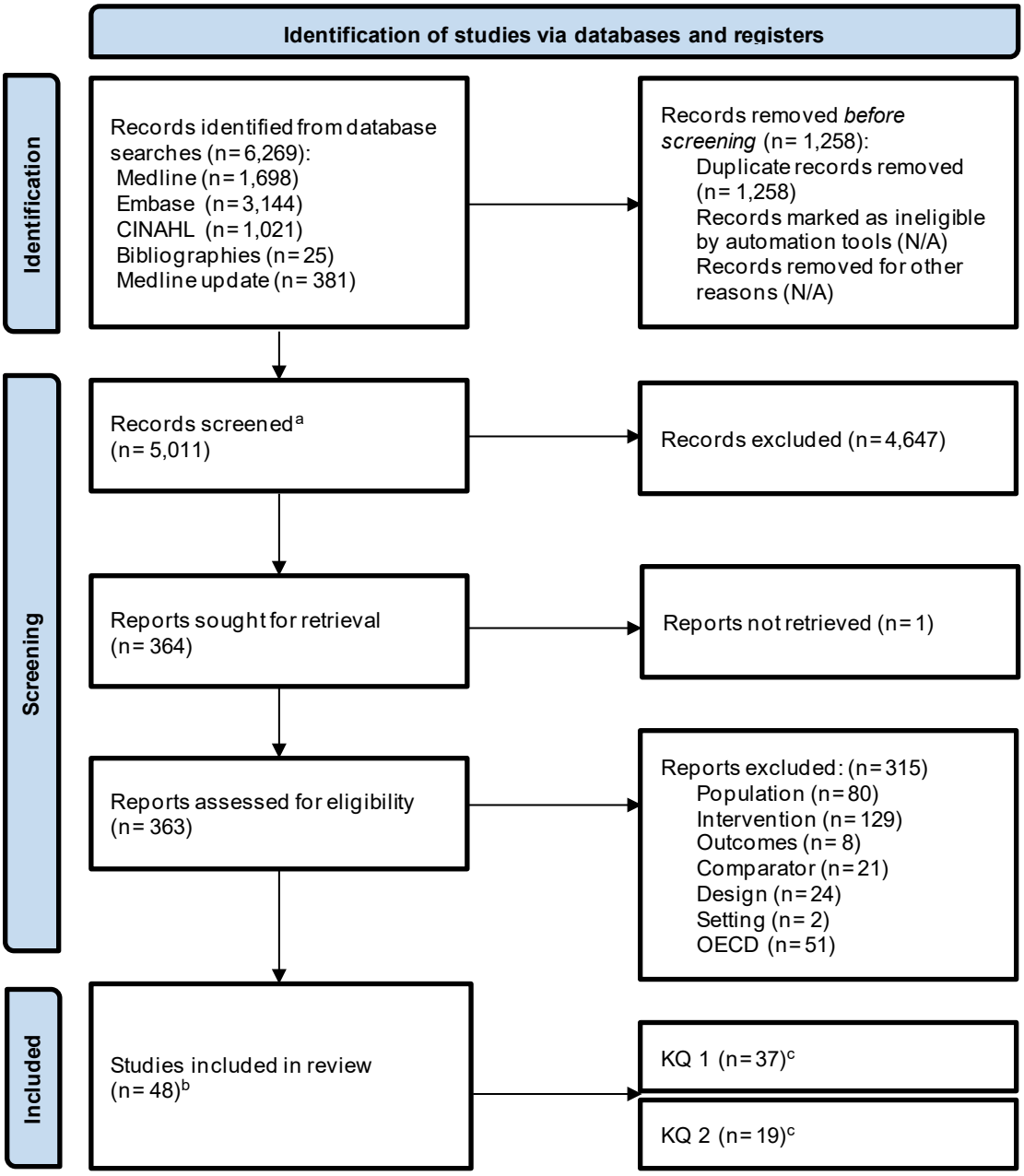
A draft version of this report was reviewed by technical experts and clinical leadership. A transcript of their comments and our responses is in Appendix F.

# RESULTS

## LITERATURE FLOW

For KQ 1 and KQ 2, we identified 5,863 studies through searches of MEDLINE® (via PubMed®), Embase (via Elsevier), and CINAHL (via EBSCO) (Figure 1). An additional 406 articles were identified after conducting a MEDLINE® update and reviewing bibliographies of relevant review articles for a total of 6,269 articles. After removing duplicates, there were a total of 5,011 articles. After applying inclusion and exclusion criteria to titles and abstracts, 364 articles remained for full-text review. Of these, 48 studies were retained for data abstraction. Of the 48 studies included, 39 were identified as unique studies. Of the 48 studies, 36 were cohort studies, 11 were cross-sectional studies, and 1 was a case-control study. Included studies were conducted in the United States (29), United Kingdom (4), South Korea (4), Canada (2), Portugal (2), Australia (2), Denmark (1), Israel (1), Japan (1), Norway (1), and Italy (1). There were 19 VA studies.

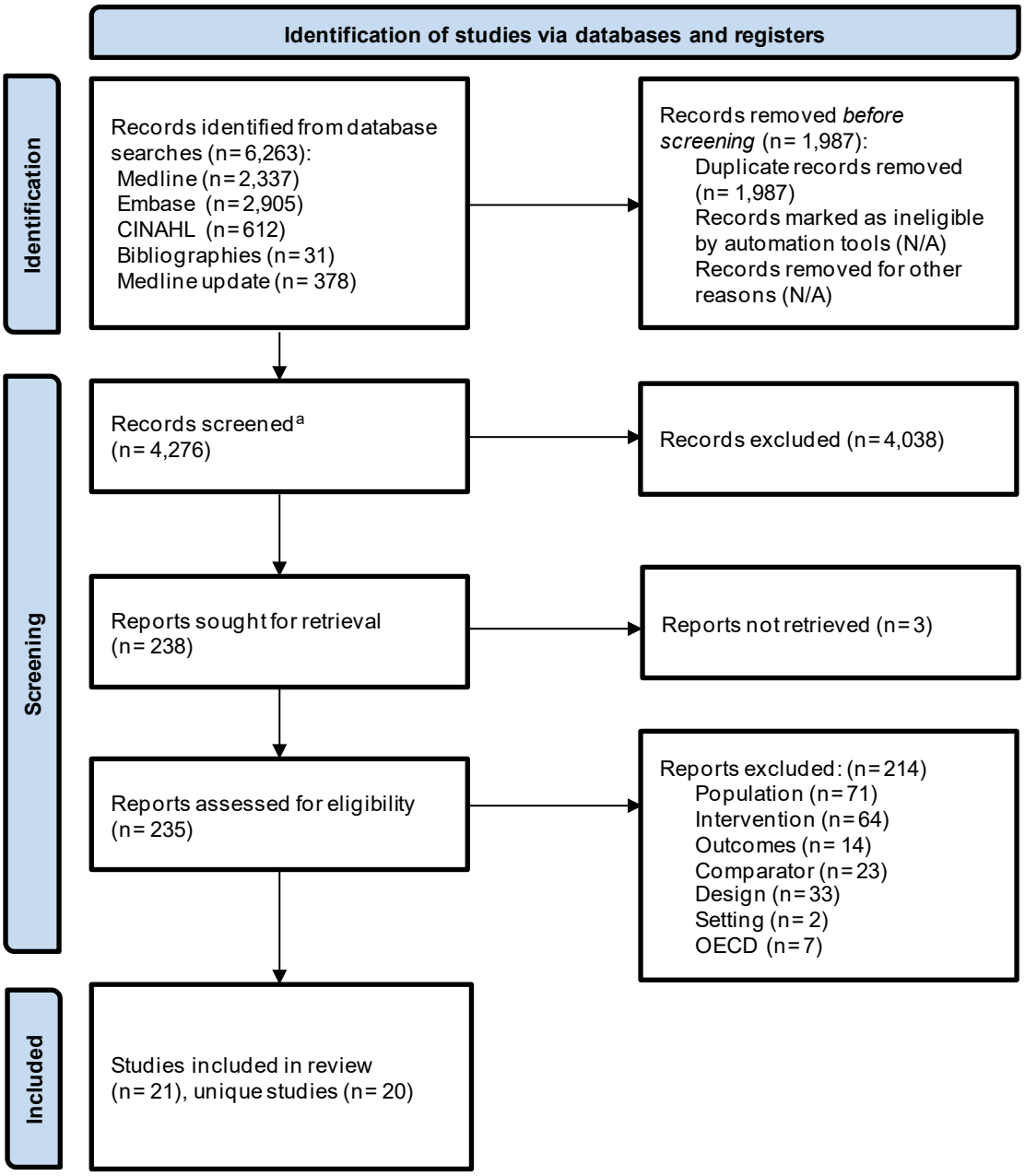
## Figure 1. Literature Flow Chart: KQ 1 and KQ 2




[a] Search results from Medline (1664), Embase (2219), CINAHL (831), manually identified (23), and Medline update (274) were combined.
[b] 10 studies use overlapping same source populations.
[c] 8 studies are included in KQ 1 and KQ 2.

For KQ 3, we identified 5,854 studies through searches of MEDLINE® (via PubMed®), Embase (via Elsevier), and CINAHL (via EBSCO) (Figure 2). An additional 409 articles were identified after conducting a MEDLINE update and reviewing bibliographies of relevant review articles, for a total of 6,263 articles. After removing duplicates, there were a total of 4,276 articles. After applying inclusion and exclusion criteria to titles and abstracts, 235 articles remained for full-text review. Of these, 21 studies were retained for data abstraction. Of the 21 studies included, 20 were identified as unique studies. Of the 20 studies, 8 were individually randomized controlled trials (RCTs), 6 were cluster-randomized trials (CRTs), 1 was a controlled before-after study, 2 were time-series studies, and 3 were nonrandomized studies. Included studies were conducted in the United States (15), Canada (3), Denmark (1), and the United Kingdom (1). None of the studies were conducted in the VA.

**Figure 2. Literature Flow Chart: KQ 3**




[a] Search results from Medline (2,331), Embase (1,349), CINAHL (297), manually identified (31), Medline update (268) were combined.

## EVIDENCE PROFILE

Table 2 shows the evidence profile of studies included in this systematic review. Appendices B and C contain detailed study characteristics for included studies.

### Table 2. Evidence Profile of Included Studies

| | KQ 1 and KQ 2 (n=48) | KQ 3 (n=20) |
|---|---|---|
| **Study design** | 36 Cohort<br>11 Cross-sectional<br>1 Case-control | 8 Randomized<br>6 Cluster-randomized<br>3 Nonrandomized<br>1 Controlled before-after<br>2 Interrupted time series |
| **Number of participants** | 12,225,464 | 114,538 |
| **Region** | 29 US<br>9 Europe<br>4 South Korea<br>2 Canada<br>4 Other | 15 USA<br>3 Canada<br>1 Denmark<br>1 UK |
| **Population** | 37 Men only<br>11 Men and women | 1 Men only<br>8 Men and women<br>11 Women only |
| **Median age (range)** | 63.5 (45 to 80.4)<br>1 study NR<br>2 studies reported age in several categories | 71.1 (51.5 to 82.0)<br>3 studies NR<br>1 study reported age in several categories |
| **Median % Male or Women (range)** | 100% Men (7% to 100%)<br>0 studies NR | 99% Women (57% to 100%)<br>0 studies NR |
| **Median % Race (range)** | 89% White (37% to 100%)<br>NR by 22 studies<br>11% Black (1% to 100%)<br>NR by 29 studies | 70% White (46% to 97%)<br>NR by 18 studies<br>14 % Black (12% to 37%)<br>NR by 18 studies |
| **Tool** | 19 FRAX<br>9 OST<br>5 QFracture<br>4 MORES<br>4 Garvan<br>6 Other | NA |
| **Intervention type** | NA | 5 Provider education<br>3 Provider and patient education<br>2 Patient navigation<br>3 Patient risk assessment<br>4 Self-referral<br>4 Provider system reminder<br>1 Patient system reminder<br>1 Clinical support tool |

| | KQ 1 and KQ 2 (n=48) | KQ 3 (n=20) |
|---|---|---|
| **Outcomes reported**[a] | 25 Osteoporosis or osteopenia via BMD<br>16 Hip fracture<br>17 Major Osteoporotic fracture<br>11 All fracture | 19 Screening<br>1 Fracture |
| **Risk of bias** | QUADAS-2[b]<br>18 At risk<br>19 Low risk<br><br>Newcastle-Ottawa[c]:<br>5 High risk<br>5 Unclear risk<br>0 Low risk<br><br>Case-control Newcastle-Ottawa[c]:<br>1 Unclear | Objective[d]:<br>2 Low risk<br>12 Unclear risk<br>2 High risk<br>2 NA<br><br>Patient-reported[e]:<br>3 Low risk<br>1 Unclear risk<br>1 High risk<br>13 NA<br><br>Interrupted time series:<br>1 Unclear risk<br>1 Low risk |

*Abbreviations*. NA=not applicable; NR=not reported
[a] Studies report more than 1 outcome type.
[b] Diagnostic test accuracy studies (29 in KQ 1 only and 8 in KQ 1 and KQ 2).
[c] Adapted Newcastle-Ottawa for cohort and case-control studies.
[d] Objective outcomes (*ie*, non–patient-reported outcomes) are not subject to a large degree of individual interpretation.
[e] Patient-reported outcomes are directly reported by the patient without interpretation of the patient's response.

# ORGANIZATION OF RESULTS

## KQ 1—Tools

This section focuses on the general population of men encompassing civilian and Veteran populations. Given the variety of screening tools identified, first we focus on the established tools described by more than 3 studies. Within each tool section, we describe the outcomes reported by each respective tool (*eg,* fracture, osteoporosis). We also describe any adaptions or slight modifications of the established tool that have been evaluated in the literature (*eg*, FRAX-A). Subsequently, we describe tools that were evaluated in patient populations at elevated risk for osteoporosis. Last, we describe the tools reported by only 1 or 2 studies.

## KQ 2—Tools and Risk Factors in Male Veterans

This section focuses on **male Veteran** populations only, including established risk assessment tools and studies identifying additional risk factors potentially salient to the Veteran population. With guidance from our operational partners, we focused on the tools and risk factors that identify Veterans at highest risk of osteoporosis or fracture. Some of these tools overlap with tools described in KQ 1, but with a specific focus on male Veterans. Thus, some studies included in KQ 1 results are also mentioned in KQ 2 and identified accordingly. Within this section, we identify the tools among general male Veteran populations as well as Veterans at elevated risk

(*eg,* HIV-positive, on androgen deprivation therapy). The independent risk factors are grouped into medical conditions only and medical conditions and exposures.

### KQ 3—Interventions

This section focuses first on systems-level interventions designed to impact providers and then on systems level interventions aimed at patients. Within each patient or provider section, studies are ordered from least intensive to most intensive intervention strategies.

## KEY QUESTION 1: Among males not identified by a history of low-trauma fracture, is there a clinical risk tool (*eg*, FRAX) that identifies patients at highest risk of osteoporosis or major osteoporotic fracture?

### Characteristics of Included Studies

We identified 37 studies that met our inclusion criteria. Included studies evaluated 18 different clinical risk tools (*eg*, FRAX, FRAX-A, FRAX-A+, OST/OSTA, QFracture, MORES, Garvan, VA-FARA, FRA HS model 1 and model 2, Korean Fracture Risk Score, Korean risk assessment, FRC, Mscore, Model I, II, and IV, Weight-based calculation). Appendix G presents a complete list of the tools and their respective components. Nine studies reported on more than 1 tool within the same population. There were 19 studies that assessed the FRAX risk assessment tool or a modified version of the FRAX risk assessment tool. Nine studies assessed the OST/OSTA. QFracture (2 different versions) was used in 5 studies. Four studies evaluated the MORES risk tool. The Garvan tool was assessed in 4 studies. The remaining tools were each evaluated in 1 study. The FRAX tool was assessed with all 3 outcomes: major osteoporotic fracture (MOF), hip fracture, and osteoporosis. The OST and the MORES, however, were only evaluated with osteoporosis as an endpoint in the identified studies. The QFracture and the Garvan tools were only compared to major osteoporotic fracture and hip fracture.

### Key Points

- We found 37 studies evaluating 18 different risk assessment tools.
- Limited evidence was identified that directly compared tools within the same population.
- Among men not identified via prior diagnosis of osteoporosis, the OST/OSTA has good discriminatory ability in predicting osteoporosis by DXA with 2, easily obtainable variables (AUC ranging from 0.632 to 0.890).
- Tools predicting osteoporosis (FRAX, MORES) reported AUC ranges from 0.596 to 0.870. High levels of heterogeneity were present.
- Tools predicting hip fracture and MOF (FRAX, QFracture, Garvan) all reported AUCs ranging from 0.609 to 0.930 for hip fracture and 0.618 to 0.810 for MOF. High levels of heterogeneity were present.
- Among men not identified via prior fracture, the FRAX risk assessment tool has better discrimination in predicting hip fracture than major osteoporotic fracture and osteoporosis diagnosis.

- Qualitatively, study location, patient age, or race did not correlate with low/moderate versus excellent discrimination.
- Limited evidence was identified for use of FRAX in special populations such as individuals with HIV and those on androgen deprivation therapy (ADT), but was generally found to perform worse among these groups.
- Limited evidence was identified using modified versions of the FRAX risk assessment tool.
- Limited evidence was identified for all other tools.

## Detailed Findings

For KQ 1, we present the detailed results ordered by clinical risk tool and, within each tool, by outcome. Details of study characteristics are in Appendix B.

### FRAX Risk Assessment Tool

The FRAX risk assessment tool (https://www.sheffield.ac.uk/FRAX/) was developed and independently validated by large international cohort studies including data from hundreds of thousands of patients. This tool incorporates multiple clinical risk factors independent of BMD (such as age, sex, weight, height, ethnicity, race, parental hip fracture, prior fracture, tobacco and alcohol use, glucocorticoid use, and rheumatoid arthritis) and secondary osteoporosis. There is the option of including femoral neck BMD to calculate a 10-year probability of hip or major osteoporotic fracture. Unlike other calculators, it does not include a history of falls.[24] Per Bone Health Osteoporosis Foundation guidelines, in the US and some other countries, the FRAX risk assessment tool is used in patients with osteopenia to determine the need for osteoporosis treatment based on treatment thresholds above 20% for major osteoporotic fracture or 3% for hip fracture.

Nineteen studies—all of which used cohort or cross-sectional designs—evaluated the FRAX risk assessment tool without BMD or a modified version of the tool in predicting 1 or more clinical outcome: major osteoporotic fracture, hip fracture, or osteoporosis (by DXA) among men. The comparator varied by study and included other fracture prediction tools, fracture rate, and/or BMD. Seven studies were conducted in the United States,[25-31] 5 in Europe (Denmark, Norway, Portugal, and 2 in the UK),[32-36] 4 in Asia (Israel, Japan, and 2 in South Korea),[37-40] 2 in Canada,[41,42] and 1 in Australia.[43] Twelve studies were among all male[25-32,34-37] populations and 2 studies were conducted specifically among male Veterans.[26,28] Three studies, which were all conducted in the United States, were among men age > 65–70 years only.[25,29,30]

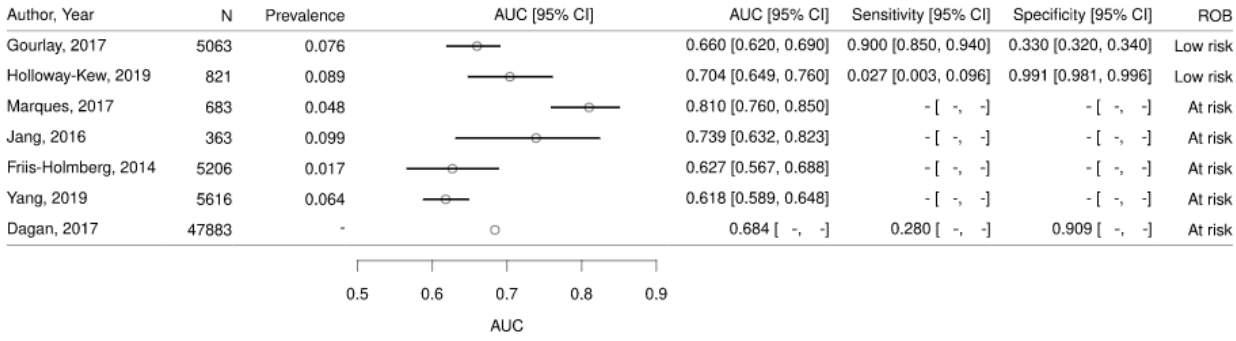
#### *Major Osteoporotic Fracture (MOF)*

Nine studies[25,30,33,36,39-43] evaluated the FRAX risk assessment tool in predicting MOF. Four studies had low risk of bias[25,30,42,43] and 5 were considered at risk[33,36,39-41] of bias. Individual study sample sizes range from 683 to 1,054,815 patients. Prevalence of major osteoporotic fracture ranged from 1.7% to 9.9% of the study populations.

Individually, most of the studies found the FRAX risk assessment tool to have poor to fair discrimination in predicting MOF (AUCs ranging from 0.618 to 0.810) (Figure 3). (Note, in general, an AUC of 0.5 is considered to mean no discrimination, 0.7 to 0.8 is acceptable

discrimination, 0.8 to 0.9 is excellent discrimination, and > 0.9 is outstanding discrimination.[44]) Six studies[30,33,36,39-41] were deemed to have sufficient conceptual homogeneity to be included in a quantitative synthesis. The 2 remaining studies[25,42] assessed populations that overlapped with 2 of the more recent studies[30,41] included in the quantitative synthesis and so were excluded. Five of the 6 studies included in the quantitative synthesis were evaluated as at risk in terms of ROB (Figures 10 and 11). Significant statistical heterogeneity was present ($I^2$ 92.5%; Q 53.4). The 2 studies not included in the forest plot had similar reported AUCs of 0.63[25] (no 95% CI reported) and 0.61[42] (95% CI 0.56 to 0.65).

**Figure 3. FRAX Tool Compared to Major Osteoporotic Fracture**

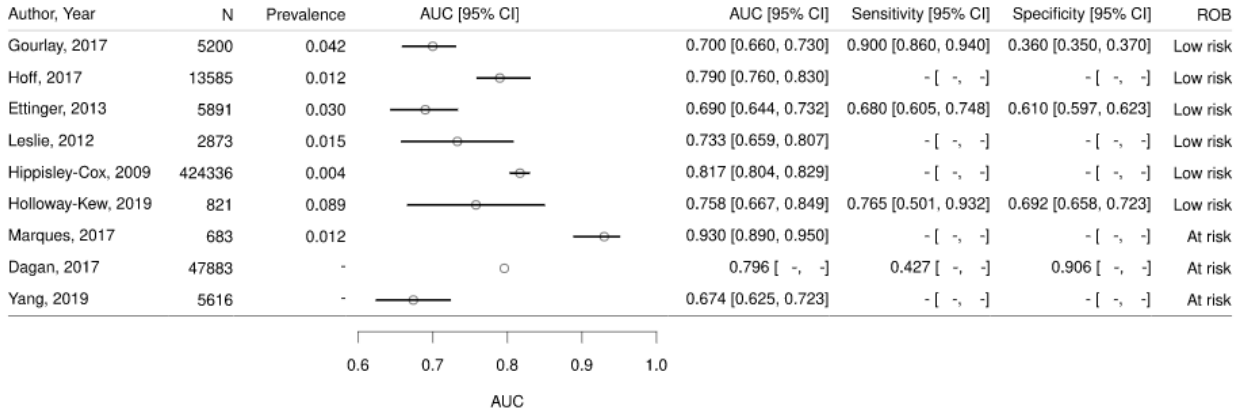



*Hip Fracture*

Nine studies[25,30,32,35,36,40-43] evaluated the FRAX risk assessment tool in predicting hip fracture rates. Six studies had low risk of bias (ROB)[25,30,32,35,42,43] and 3 were at risk.[36,40,41] Individual study sample sizes range from 683 to 424,336 patients. Prevalence of hip fracture ranged from 0.4% to 8.9% (unable to calculate in 2 studies[40,41]).

Individually, most studies found the FRAX risk assessment tool to have better discrimination in predicting hip fracture (AUC ranging from 0.67 to 0.93) than major osteoporotic fracture (Figure 4) Of the 5 studies[30,36,40,41,43] that assessed both outcomes, discrimination was consistently better for hip fracture compared to major osteoporotic fracture.

**Figure 4. FRAX Tool Compared to Hip Fracture**



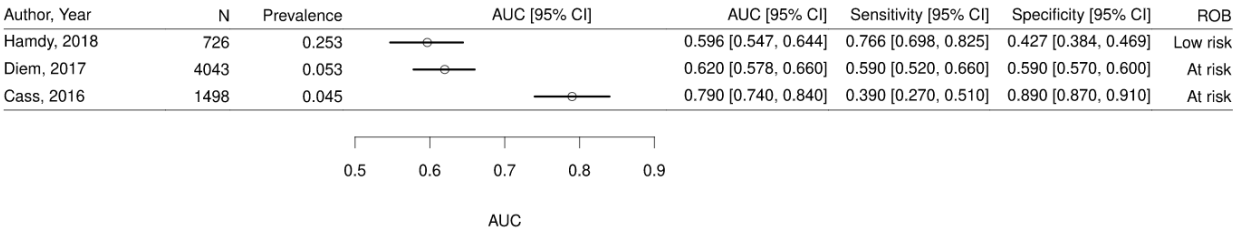
### *Osteoporosis*

Four studies[27,29,31,37] evaluated the FRAX risk assessment tool in predicting diagnosis of osteoporosis by DXA. One study[31] had low ROB and 3 were at risk[27,29,37]of bias. Individual study sample sizes range from 520 to 4043 patients. The prevalence of osteoporosis across studies ranged from 5.3% to 25.3%. The prevalence of osteoporosis in 1 study population was not reported.[37]

Individually, most studies found the FRAX risk assessment tool to have poor discrimination in predicting osteoporosis by DXA (AUCs ranging from 0.596 to 0.790). Three of the studies[27,29,31] could be included in quantitative synthesis (Figure 5). One study[31] had higher prevalence of 25.3% compared to the other 2 included studies[27,29] (4.5% and 5.3%). Significant statistical heterogeneity was present ($I^2$ 94.5; Q 36.3). There was variation across studies in what reference populations were used to calculate T-scores *(eg,* NHANES III female, NHANES III male, Young adult mean)

The 1 study[37] not included in the forest plot reported results for men < 75 years of age and men ≥ 75 years separately. AUCs for these groups were 0.63 (95% CI 0.49 to 0.77) for men < 75 years and 0.67 (95% CI 0.59 to 0.75) for men ≥ 75 years.

**Figure 5. FRAX Tool Compared to Osteoporosis**




### *FRAX Risk Assessment Tool with Modifications*

Due to limitations identified in predicting fracture risk in certain individuals, modifications to the FRAX risk assessment tool have been proposed. These include modification when the spine BMD is disproportionately lower than the hip BMD, in individuals with diabetes,[45] and in individuals on high doses of glucocorticoids,[46] among others.[47] Four studies[26,28,34,41] included in this review evaluated a modified FRAX risk assessment tool. The 2 studies[26,34] in men with HIV discussed above also calculated FRAX risk assessment scores using HIV as a secondary cause of osteoporosis. In the study conducted at the VA,[26] the modified FRAX scores changed the observed/expected ratio for MOF for men with HIV 1.62 to 1.20 (95% CI 1.08 to 1.34) and from 4.52 to 2.66 (95% CI 2.17 to 3.26) for hip fracture. For MOF, none of the men met the > 20% threshold endorsed by the NOF, so the age-specific thresholds (6.3% to 13.4% in 50–70-year-olds) endorsed by the European osteoporosis societies were utilized. For hip fracture, the > 3% hip fracture threshold endorsed by the NOF was utilized. Among men with HIV using these thresholds, the sensitivity for MOF was 6.4% and hip fracture was 3.2%. The sensitivity among men without HIV was 2.6% for MOF and 0% for hip fracture. In the second study conducted in the UK,[34] utilizing HIV as a secondary cause of osteoporosis changed the sensitivity of the modified FRAX risk assessment tool from 23% to 31% and specificity from 88% to 75%. One study that investigated a modified FRAX score, coined "e-FRAX," that involved risk factor

ascertainment via EHR.[28] In the e-FRAX tool, parental hip fracture was the only risk factor assumed to be absent since documentation of this is known to be poor in the EHR. FRAX had better accuracy than e-FRAX (AUC 0.72 [95% CI 0.67 to 0.78] vs 0.65 [95% CI 0.59 to 0.71]) for predicting osteoporosis in this study. The fourth study[41] was conducted in Canada and included 5615 men (9.2% total included). This study was at risk for bias. Four modified FRAX risk assessment models were developed based on different combinations of clinical risk factors included in FRAX and administrative data. AUCs for these modified FRAX scores for MOF were 0.584 (95% CI 0.553 to 0.615) for FRAX (age-sex), 0.624 (95% CI 0.594 to 0.654) for FRAX (age-sex-fracture), 0.616 (95% CI 0.586 to 0.646) for FRAX A (*ie*, FRAX, except for BMD, BMI, and parental hip fracture), and 0.648 (95% CI 0.619 to 0.677) for FRAX A+ (*ie,* FRAX A, plus a comorbidity score, number of hospitalizations in the 3 years prior, depression, and dementia). AUCs for the modified FRAX risk assessment scores for hip fracture were 0.663 (95% CI 0.612 to 0.714) for FRAX (age-sex), 0.657 (95% CI 0.605 to 0.709) for FRAX (age-sex-fracture), 0.648 (95% CI 0.598 to 0.698) for FRAX A, and 0.676 (95% CI 0.626 to 0.727) for FRAX A+.

### *FRAX Risk Assessment Tool in Patients at Higher Risk*

Three studies[26,34,48] evaluated the FRAX risk assessment tool in special populations; in men with HIV in 2 of the studies[26,34] and in men undergoing androgen deprivation therapy (ADT) in 1.[48] One of the studies[26] focusing on men with HIV was conducted at the VA and included 24,451 Veterans. The observed/expected ratio of MOF for men with HIV was 1.62 (95% CI 1.45 to 1.81) versus for men without HIV 1.29 (95% CI 1.19 to 1.40) with a p-value of 0.03 for the difference between men with and without HIV. For hip fracture, the observed/expected ratio was 4.52 (95% CI 3.68 to 5.53) for men with HIV versus men without HIV 3.56 (95% CI 3.03 to 4.18). This study was at risk of bias. The second study[34] of men with HIV was conducted in the UK with younger men aged 38 to 51 years. This study was low risk of bias and included 168 men. It utilized a threshold of ≥ 7.5% for the FRAX risk assessment tool in predicting osteoporosis (T-score ≤ -2.5) or low bone density for age (Z-score ≤ -2.0) by DXA. With this threshold, the sensitivity was 23% and specificity was 88%. The third special population study[48] evaluated FRAX prediction of MOF or hip fracture in men undergoing androgen deprivation therapy. This low ROB study included 115 Veterans with a mean age of 77 years. Based on the NOF thresholds of risk of major osteoporotic fracture ≥ 20% or risk of hip fracture ≥ 3%, FRAX without BMD recommended treatment in 54% of Veterans. If treatment was based on T-score by DXA alone, 35% would be recommended for treatment due to T-score ≤ -2.5, 54% for T-score ≤ -2.0, and 69% for T-score ≤ -1.5. Note, no AUCs were reported in the studies of special populations.

## Osteoporosis Self-assessment Tool (OST) and Osteoporosis Self-assessment Tool for Asians (OSTA)

### *Osteoporosis*

The OSTA was first established to predict osteoporosis among Asian women and is calculated using the following formula: OSTA = (weight in kg – age in years) x 0.2. The result is rounded down to the nearest whole number and categorized as low risk (> -1), moderate risk (-1 to -4), or high risk (< -4). Subsequently, the OST was validated in non-Asians using the same calculation and different cutoff values. A cutoff of OST < 2 was predictive of osteoporosis by DXA in Caucasian women.[49]

Seven studies evaluated the OST, including 5 prospective cohort studies[29,50-53] and 2 cross-sectional studies.[28,54] One cross-sectional study[55] utilized the OSTA. Using the QUADAS, 1 of the OST studies[28] and the 1 OSTA study[55] were determined to have a low risk of bias; the remaining studies were considered at risk of bias.

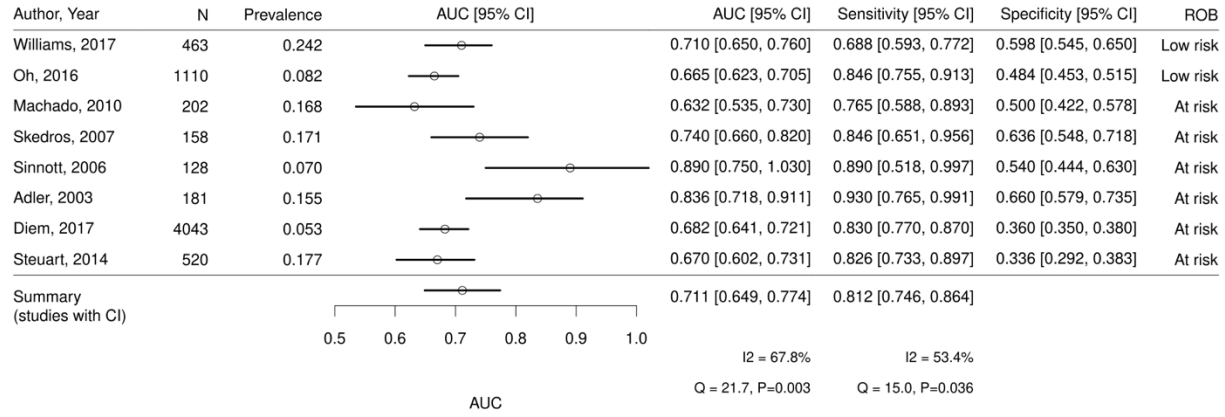
The eight included studies comprised 6,075 participants in total. The mean/median age of participants ranged from 63.5 to 80.4 years. Six of the studies were mostly white (68.5% to 100%), while 1 study utilized data only collected from African American male Veterans (n = 128).[51] Two other studies were also conducted at a VA Medical Center and included a total of 644 Veterans.[28,54] The OSTA study was composed of Korean men.[55]

Each of the included studies aimed to determine the optimal cutoff for the OST in predicting osteoporosis by DXA for the selected population. Three of the studies utilized whole-body DXA scanning, while 5 of the studies based DXA score on scans of specific skeletal locations (various combinations of the spine, femoral neck, radius, and hip).[28,29,50,52,55] There was variation across studies in whether a female or male reference population was used to calculate T-scores, and this was not specified in some studies.[28,52] The prevalence of osteoporosis in the studies ranged from 5.3%[29] to 24.2 %.[28] The optimal cutoff of OST ranged from 0.99 to 6 with AUCs for these cutoffs ranging from 0.632 to 0.836. In the study of the OSTA, the predefined cutoff of 0 had a sensitivity of 86.2%, specificity of 49.7%, and AUC of 0.680.

Overall results for this synthesis are presented in Figure 6. A meta-analysis of the 8 studies was performed yielding a summary AUC of 0.71 (95% CI 0.659 to 0.77). A Cochrane's Q of 21.7 ($p=0.003$) and $I^2$ of 67.8% suggest that there is substantial heterogeneity and variability between the studies contributing to this estimate. This score indicates that the OST possesses acceptable diagnostic accuracy. Sensitivity of the OST was also presented as pooled values across the included studies 81.2% (95% CI 74.6 to 86.4). Specificity estimates were not pooled given substantial heterogeneity ($I^2$ 96.4%, Q 195.3, $p<0.001$).

Two of the studies presented subgroup analyses for age and race/ethnicity. Adler et al 2003 obtained similar results after grouping subjects by age (grouped by decades) and race/ethnicity.[54] However, Richards identified a different optimal cutoff for OST in predicting osteoporosis among Caucasian men (OST cutoff ≤ 5, sensitivity: 75.4%, specificity: 41.4%) versus African American men (OST cutoff ≤ 6, sensitivity: 70.0%, specificity: 36.4%).[53] This same author suggested that age may impact the predictability of the OST, with subjects > 65 and an OST cutoff ≤ 2 yielding a sensitivity of 80.0% and a specificity of 52.8% versus subjects ≤ 65 and an OST cutoff ≤7 yielding a sensitivity of 76.2% and a specificity of 39.5%.

**Figure 6. OST Tool Compared to Osteoporosis**




### *OST Risk Assessment Tool in Patients at Higher Risk*

The 1 study not included in the forest plot reported the performance of the OST tool among Veterans with rheumatoid arthritis.[56] This study did not report an overall AUC. However, at a score of 4 or less, the tool correctly identified 78% of the population with osteoporosis and correctly identified 45% as not having osteoporosis compared to a DXA (*ie,* sensitivity 78%, specificity 45%).

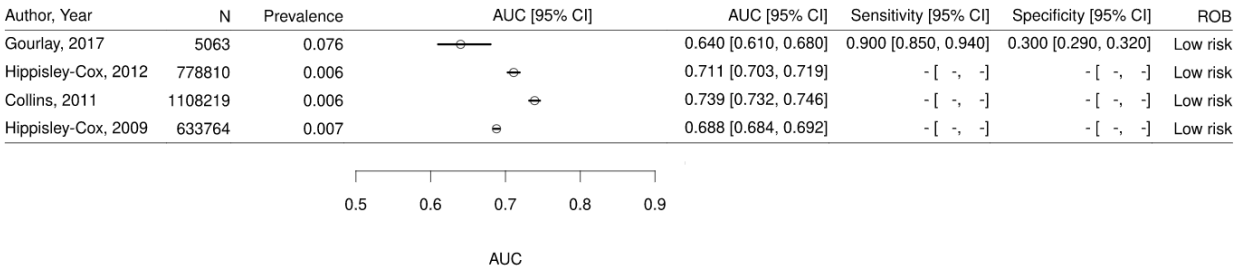
## QFracture

The QFracture Tool was developed to estimate the risk of major osteoporotic fracture (hip, spine, wrist, or shoulder) or hip fracture over the next 10 years. It was developed and validated utilizing a prospective cohort of over 2 million men (49.8% men) and women aged 30–85 years from multiple UK primary care practices.[32] It includes many of the same clinical risk factors for osteoporosis as the FRAX risk assessment tool with the addition of fall history and 10 others but does not include BMD. We identified 5 prospective/retrospective cohort studies evaluating the QFracture tool in predicting major osteoporotic fracture and hip fracture.[30,32,40,57,58] The studies compared the QFracture tool to other risk assessment tools and/or to fracture rates. Three of the studies utilized UK cohorts.[32,57,58] One study was centered in the US[30] and another in Israel.[40] Over 6 million men were included in these studies and men aged 30 to 100 years were included. Four of the 5 studies evaluating the QFracture tool had a low risk of bias[30,32,57,58] and 1 was at risk of bias.[40]

### *Major Osteoporotic Fracture*

These studies all evaluated the QFracture tool in predicting major osteoporotic fracture. The 10-year prevalence of major osteoporotic fracture ranged from 0.6% to 7.6% across the studies.[30,32,57,58] The QFracture has poor to fair discrimination among the 4 studies evaluating a 10-year prediction of major osteoporotic fracture (AUC 0.640 to 0.739)[30,32,57,58] (Figure 7). Significant statistical heterogeneity was present ($I^2$ 98.2%; Q 166.3). The remaining study found poor discrimination of the QFracture tool in predicting 5-year major osteoporotic fracture risk (AUC 0.686).[40]

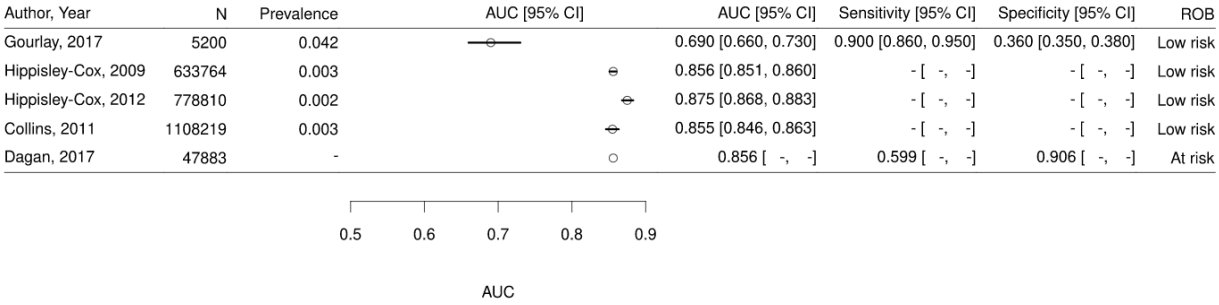
**Figure 7. QFracture Tool Compared to Major Osteoporotic Fracture**




### *Hip Fracture*

Each of the 5 studies also evaluated the QFracture tool in predicting hip fracture. The 10-year hip fracture prevalence ranged from 0.2% to 4.2%.[30,32,57,58] Overall, the QFracture had better discrimination for 10-year prediction of hip fracture compared to major osteoporotic fracture with AUC ranging from 0.690 to 0.875 (Figure 8). Significant statistical heterogeneity was present ($I^2$ 97.3%; Q 109.4). The study of 5-year prediction of hip fracture also found excellent discrimination with an AUC of 0.856.[40]

**Figure 8. QFracture Tool Compared to Hip Fracture**




## Male Osteoporosis Risk Estimation Score (MORES)

The MORES was developed to predict osteoporosis at the total hip specifically among men ≥ 50 years and utilizes age, weight in kg, and history of COPD.[59] The established cutoff is a score ≥ 6 (out of 20) in predicting osteoporosis (sensitivity 93%, specificity 59%, AUC 0.832). Data from 2,995 men were used to develop and validate the algorithm. Shepherd et al later assessed the MORES in predicting osteoporosis at the vertebra or any site.[60] The original cross-sectional study of the MORES development and validation[59] and the secondary study assessing the MORES for these additional sites are included in this review. Two additional cross-sectional studies evaluating the MORES in predicting osteoporosis are also included in this review.[27,61] BMD obtained by DXA served as the gold standard for the diagnosis of osteoporosis in each of the studies for MORES. T-scores were calculated according to varying reference populations across the studies. All 4 studies were conducted in the US and included predominantly white males (76%–88.5%) with mean ages of 63 to 70.2 years. A total of 7,823 men were included in these 4 studies.

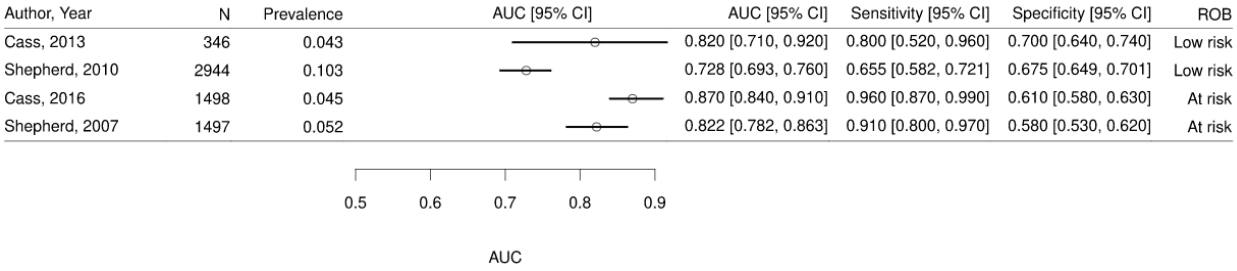
### *Osteoporosis*

The included studies evaluated the MORES in predicting osteoporosis by DXA. Two studies were low risk of bias[60,61] and 2 were at risk.[27,59] The prevalence of osteoporosis across studies

ranged from 4.3% to 12.4%. Three of the 4 studies assessed for osteoporosis at the hip.[27,59,61] One study assessed for osteoporosis at any site (hip or vertebral) and vertebral only. Data from all 4 studies (excluding the validation cohort in 1 study[59] and the vertebral osteoporosis only outcome in another study[60]) were combined in a meta-analysis. Individually, these studies found the MORES to have fair to excellent discrimination in predicting osteoporosis (AUC 0.728 to 0.870) (Figure 9). Significant statistical heterogeneity was present ($I^2$ 91.3%; Q 34.4). The study with the lowest AUC[60] in the plot was the study evaluating the prediction of osteoporosis at any site (hip or vertebral) versus osteoporosis at the hip only among the 3 other studies.

Among the data not included in the plot, the data from the validation cohort of 1 study[59] had a similar reported AUC of 0.842 (95% CI 0.811 to 0.873) whereas the data from the vertebral osteoporosis only outcome had poor discrimination with AUC of 0.657.[60] None of the identified studies evaluated the MORES tool for MOF or hip fracture.

**Figure 9. MORES Tool Compared to Osteoporosis**




The Garvan model was developed by the Dubbo Osteoporosis Epidemiology Study (DOES) that began in 1989. It predicts the 10-year absolute risk of hip fracture or major osteoporotic fracture. They compared over 50 risk factors and determined that 5 risk factors[62] accounted for the greatest portion of the variance of risk. The 5 risk factors are: age, bone mineral density, body weight, history of prior fracture after the age of 50, and any falls during the past 12 months.[62] Four studies[30,40,43,63] evaluated the Garvan model. Three were both low risk of bias cohort studies; 1 was conducted in Israel,[40] 1 in Australia,[43] and the other in the United States, called MrOS.[30] The MrOS had 5,200 men aged 65 years and older, and the study conducted in Israel had 1,054,815 total participants, of which 478,825 were men ages 50-90. The Nguyen et al cohort study enrolled 2,216 individuals (858 were men) over the age of 60 from the DOES study.[63] They developed a nomogram to predict a 5-year and 10-year absolute fracture risk.[63] This at risk of bias study was conducted in Australia and evaluated models with the same risk factors, later known as the Garvan model. We included model II in the review since it did not contain BMD.

### *Major Osteoporotic Fracture*

Two studies reported Garvan predicting the risk of MOF.[30,63] One study had an AUC of 0.66 (95% CI 0.62 to 0.70) reporting p = 0.4517.[30] The Garvan nomogram study had an AUC of 0.739 with a SE of 0.024 and reporting p = 0.0240 for predicting a 5- or 10-year risk of major osteoporotic fracture in men showing fair discrimination.[63]

### *Hip Fracture*

Three studies[30,40,43] evaluated the Garvan tool and found fair to good discrimination in predicting the 10-year risk of hip fracture. Gourlay et al reported fair discrimination in predicting a 10-year risk of fracture with an AUC of 0.71 (95% CI 0.67 to 0.74)[30] and Dagan et al found good discrimination in predicting 10-year risk of fracture with an AUC of 0.765.[40] Holloway-Kew et al reported a similar AUC of 0.773 (95% CI 0.691 to 0.855).[43] In the tools assessed, the AUC discrimination improved when predicting hip fracture compared to the tools predicting a major osteoporotic fracture.

## Other Clinical Risk Assessment Tools

In addition to the previously described tools, we identified 6 studies[28,55,64-67] describing 6 additional tools including the FRActure Health Search (FRA-HS) score, Fracture Risk Calculator (FRC), KORAM-M, VA-FARA, Mscore, and the Korean Fracture Risk Score (KFRS). These tools are less common and are described by 2 or fewer included studies.

First, the FRA-HS was developed and validated by Francesco et al for use in primary care settings in Italy. It is a FRAX-based tool[64] to calculate the 10-year predicted risk of osteoporotic fracture. The FRA-HS consists of the risk factors: BMI, sex, age, long-term use of corticosteroids, alcohol abuse or alcohol-related diseases, smoking status, rheumatoid arthritis, history of osteoporotic fractures, and other causes of secondary osteoporosis. It was a cohort study with a low risk of bias rating that included 407,771 total participants, of which 183,308 were men 40 years or older.

Second, the Fracture Risk Calculator (FRC) was based on the National Osteoporosis Foundation's selection of key risk factors.[66] The FRC provides a 10-year risk estimate of both hip fracture and major osteoporotic fracture (*ie,* hip, clinical spine, forearm, shoulder). Ettinger et al[66] compared the FRC with and without BMD to see if BMD would affect the performance of the tool. The FRC consists of the risk factors: age, sex, race/ethnicity, BMI, BMD, history of fracture, parental history of hip fracture, smoking and alcohol consumption, use of corticosteroids, the prevalence of rheumatoid arthritis, and secondary osteoporosis. This low risk of bias cohort study was conducted in the United States with 5,893 men aged 65 years and older.

Third, the KORAM-M was developed and validated by Oh et al to identify Korean men at high risk of developing osteoporosis based on the nationwide dataset.[55] They evaluated 3 models: Model 1 consisted of age and weight; Model 2 age, weight, and health behavior; Model 3 age, weight, exercise, and blood tests. This low risk of bias cross-sectional study was conducted in South Korea and selected 2,450 men 50 years and older.

Fourth, Williams et al compared the performance of 4 tools: FRAX, e-FRAX, OST, and the VA-FARA.[28] The VA-FARA (Veterans Affairs Fracture Absolute Risk Assessment tool) was designed to identify the risk of fracture as correlated with osteoporosis. The risk factors the VA-FARA consisted of were prior fracture, age > 80, underweight, malnutrition, opioid exposure, proton-pump inhibitor use, depression diagnosis, stroke, seizure disorder, alcohol abuse disorder, fall risk, and clinic visits in prior year. This low risk of bias cross-sectional study was conducted in the United States from the Salt Lake City VA and consisted of 463 men 70 years and older.

Fifth, the Mscore[67] was developed by Zimering et al and modeled after SCORE for male Veterans. They compared test characteristics in Caucasians and African Americans. It is a weighted risk index for osteoporosis consisting of the variables age, weight, gastrectomy, emphysema, and prior fractures. This at risk of bias cohort study was conducted in the United States and analyzed a total of 970 men ages 40 years and older. They had a development cohort (n = 639), a validation cohort (n = 197), and an African American cohort (n = 134).

Finally, the Korean Fracture Risk Score (KFRS) was developed and validated by Kim et al as an Asian-specific prediction model.[68] The score predicts a 7-year risk of osteoporotic fracture and consists of the variables: age, body mass index, recent fragility fracture, current smoking status, high alcohol intake, lack of regular exercise, recent use of an oral glucocorticoid, rheumatoid arthritis, and other causes of secondary osteoporosis. This low risk of bias cohort study included 718,306 total participants, of which 370,255 were men ages 50–90 who were enrolled in the Korean NHIS database. They used both a modeling cohort (n = 185,127) and a validation cohort (n = 185,128).

### *Major Osteoporotic Fracture*

There were 4 studies[28,64,66,68] evaluating how these emerging tools predict MOF; the overall discrimination was found to be poor to fair for all *6* studies. The FRA-HS tool had an AUC of 0.49 (95% CI 0.48 to 0.50) for predicting a 10-year risk of major osteoporotic fracture in men showing poor discrimination.[64] The FRC had an AUC of 0.66, the KFRS had an AUC of 0.68, and the VA-FARA had an AUC of 0.64 (95% CI 0.58 to 0.70).

### *Hip Fracture*

There were 2 studies[64,66] that evaluated these tools on the prediction of hip fracture rates; the overall discrimination was found to be poor to good for all 5 studies. The FRA-HS tool was found to have poor discrimination for risk of hip fracture in men with an AUC of 0.66 and a 95% CI [0.64–0.68].[64] The FRC has fair discrimination for risk of hip fracture in men with an AUC of 0.71, a sensitivity of 0.74, and a specificity of 0.57.[66]

### *Osteoporosis*

There were 2 studies that evaluated tools on prediction risk of osteoporosis.[55,67] The Mscore was found to have excellent discrimination for the risk of osteoporosis in men with an AUC of 0.84 and a 95% CI 0.74 to 0.95. The 5-variable Mscore was not tested in the African American cohort. The KORAM-M tool was found to have poor discrimination in identifying the risk of osteoporosis; when using a cutoff ≤ -9 resulted in an AUC of 0.638 (SE 0.019), an AUC of 0.618 (SE 0.020) at a cutoff ≤ -10, and AUC of 0.642 (SE 0.018) at a cutoff of ≤ -12.

Overall, for the emerging tools, the discrimination in AUC was relatively higher when predicting hip fracture rates. Despite the Mscore having excellent discrimination when predicting osteoporosis, the validation cohort had the smallest number of participants (n = 197).

## QUALITY OF EVIDENCE FOR KEY QUESTION 1

The ROB was judged at risk for 18 studies and low risk for 18 studies. Patterns that led to judgments of at risk included: 1) patient selection (n = 11), 2) interpretation of index test (*eg,*

FRAX, OST) (n = 13), 3) interpretation of the reference standard (*eg,* DXA) (n = 9), and 4) patient flow and timing (n = 12). ROB ratings and assessments for each study are shown in Figures 10 and 11.

## Figure 10. Risk of Bias Ratings for the Included KQ 1 Studies

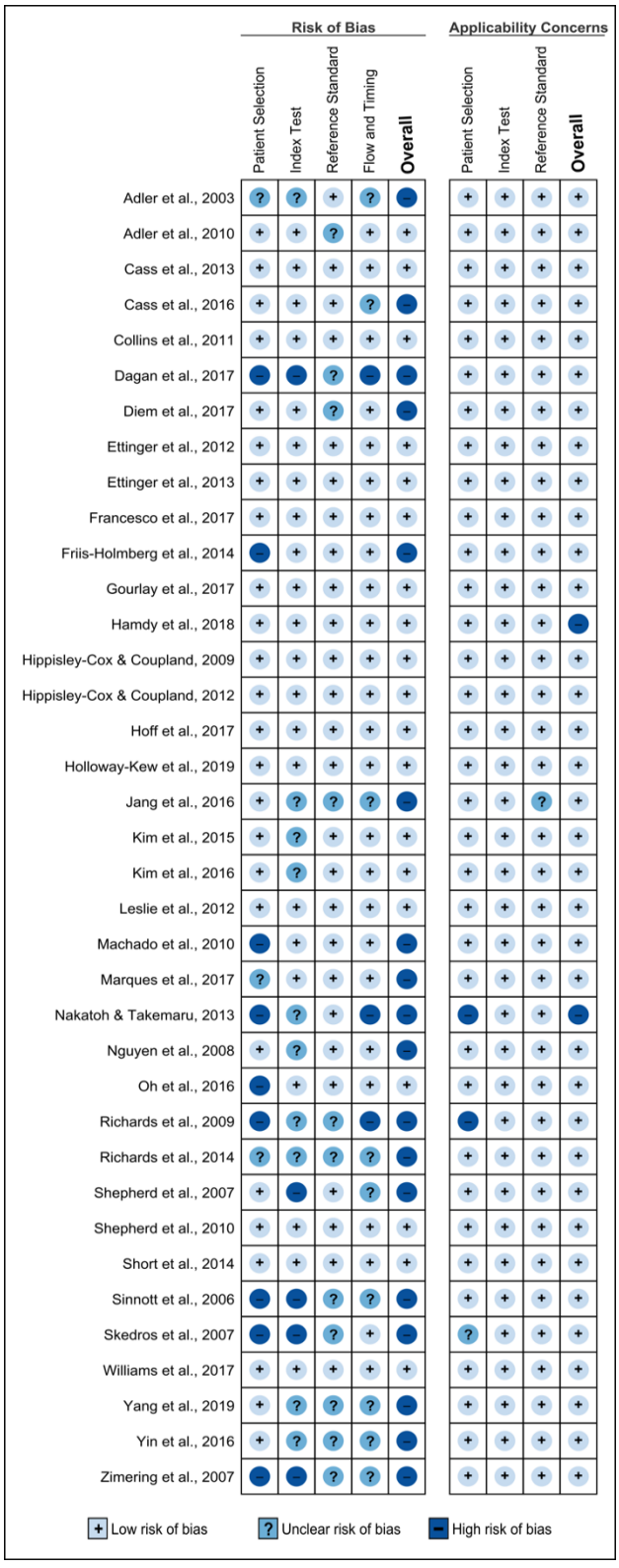
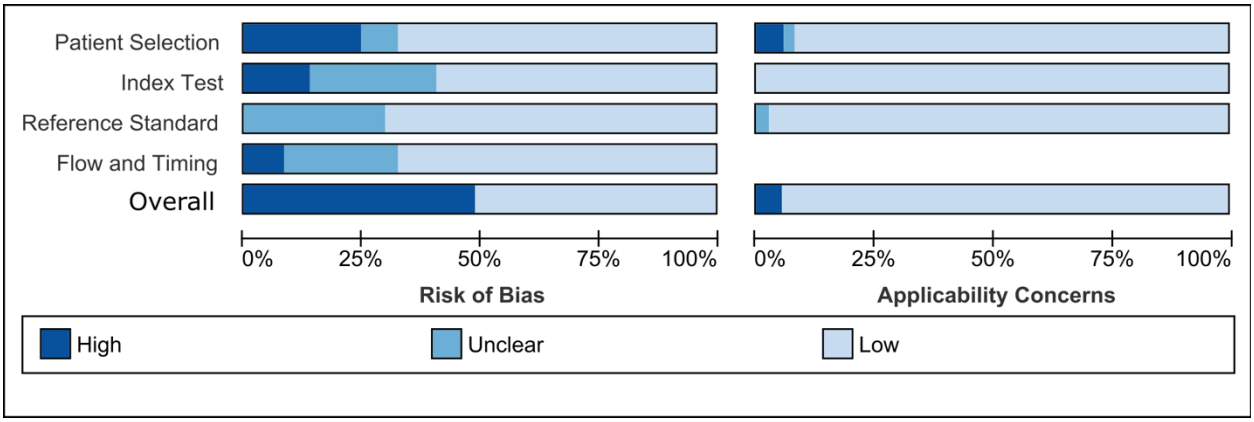


**Figure 11. Risk of Bias Assessment Across Included KQ 1 Studies**




## KEY QUESTION 2: Among male Veterans not identified by a history of low-trauma fracture, is there a tool or combination of risk factors that identify patients at highest risk of osteoporosis or major osteoporotic fracture?

### Characteristics of Included Studies

To address this question, we evaluated the subset of studies conducted specifically in **male Veterans** not identified by a history of low-trauma fracture. These studies examined individual risk factors or risk assessment tools and their association with osteoporosis or osteopenia, defined by T-scores on DXA, and fracture defined by diagnosis codes. Some of the studies described in this section have been previously discussed in KQ 1 results above, but here are revisited within the context of Veteran-specific analyses. Specifically, 8 studies appear both in KQ 1 and KQ 2.[26,28,48,51,53,54,56,67]

Eight studies[26,28,48,51,53,54,56,67](n = 26,469) examined risk assessment tools, and of these, three[26,48,56] (n = 24,848) were conducted in populations of special interest where fracture risk is considered higher than the general population. Twelve studies[51,69-79] (n = 585,400) assessed individual risk factors for low BMD and/or fracture and risk factors were broadly categorized as conditions or exposures (*eg,* smoking, alcohol, medications). Mean age varied across studies, from 55.6 to 80.4 years in studies examining risk tools, and 46 to 76 years in those evaluating individual risk factors. Race/ethnicity of studied populations varied widely, as illustrated in Table 2.

Overall, there was substantial conceptual heterogeneity across studies in terms of: a) which risk factors were utilized in tools; b) how risk factors were obtained (*ie*, patient report vs obtained from EHR); c) cutoffs used for the same risk prediction tools (*ie,* non-guideline recommended cutoffs for OST score); and d) how outcomes were defined (*eg,* diagnosis codes, T-scores). Appendix B summarizes study characteristics. Additionally, when BMD measurements by DXA were utilized to define osteoporosis, the reference population for T-scores varied across studies, and in some cases was not reported (3 studies[28,76,80]).

## Key Points

- Many studies report tools predicting osteoporosis and fracture (n = 8) as well as independent risk factors (n = 12) among male Veteran populations.
- Tools perform similarly **among male Veterans** compared to other male populations:
  - OST/OSTA predicted osteoporosis similarly among Veterans (AUC 0.670 to 0.890) as among general populations (AUC 0.632 to 0.740).
  - FRAX had an AUC of 0.72 (95% CI 0.67 to 0.78) for predicting osteoporosis in 1 Veteran study, and AUC 0.596 to 0.870 in general populations.
- Among an average-risk male Veteran population, tools using combinations of risk factors had moderate discriminant validity to predict osteoporosis and/or fracture:
  - FRAX and OST were the most common tools used to assess combinations of risk factors for predicting osteoporosis and/or fracture among male Veterans.
  - A single study suggests the Mscore may perform well for predicting osteoporosis among male Veterans.
- Tools using combinations of risk factors to predict osteoporosis and/or fracture in **Veterans at high risk for fracture** had low/moderate discriminant validity:
  - Among male Veterans, FRAX appears to underestimate risk of fracture in HIV and HCV infection, as well as in those treated with ADT.
  - Compared to its performance in average risk male Veteran populations, OST appears to perform sub-optimally for predicting osteoporosis in male Veterans with rheumatoid arthritis.
- Among male Veterans, we identified limited evidence supporting individual risk factors for osteoporosis and/or fracture. All included studies risk factor studies were at high or unclear risk of bias.

## Detailed Findings

### Risk Assessment Tools (Using Combinations of Risk Factors) to Predict Osteoporosis and/or Fracture in Male Veterans

We identified eight[26,28,48,51,53,54,56,67] studies (n = 26,469) that assessed combinations of risk factors for osteoporosis and/or fracture in the form of risk assessment tools specifically among male Veterans (note: all of these studies were also reported above in KQ 1). The most common tools examined were the FRAX (or modified versions of FRAX) and OST. Other tools included VA-FARA and Mscore; however, these were examined in only 1 study each.[28,67] Three studies[26,48,56] (n = 24,848) were conducted on populations of special interest who may be at heightened risk of osteoporosis and fracture: (1) human immunodeficiency virus (HIV) and hepatitis C virus (HCV) infection (FRAX)[26]; (2) history of androgen deprivation therapy for localized prostate cancer (FRAX)[48]; and (3) rheumatoid arthritis (OST)[56]. Only 1 study (n = 463) conducted a direct head-to-head comparison of different risk assessment tools,[28] and this study compared FRAX, OST, and VA FARA using electronic health record (EHR) data from an average risk male Veteran population.

### *FRAX and Modified FRAX Scores*

*Performance in Average Risk Veterans*

We found only 1 study that investigated FRAX and a modified FRAX in predicting osteoporosis (T-score ≤ –2.5 on DXA) among 463 male Veterans with no known risk factors for accelerated bone loss.[28] Their approach to the standard FRAX calculation involved obtaining risk factors through a patient questionnaire. They also examined a modified FRAX score, coined "e-FRAX," that involved risk factor ascertainment via EHR review. In the e-FRAX tool, parental hip fracture was the only risk factor assumed to be absent since documentation of this is known to be poor in the EHR. FRAX had better accuracy than e-FRAX (AUC 0.72 [95% CI 0.67 to 0.78] vs 0.65 [95% CI 0.59 to 0.71]) for predicting osteoporosis in this study, which may be attributable to a more accurate representation of risk factors when obtained by patient self-report versus EHR extraction. Given that few studies evaluating FRAX in Veteran populations used similar estimates (*ie*, AUC, sensitivity, specificity), our ability to draw comparisons to non-Veteran populations is limited. However, they seem to perform similarly (AUC 0.596 to 0.870 in general populations). The majority of Veterans in this study were Caucasian (94.2%), so additional data would be needed to understand the performance of FRAX ad e-FRAX in a racially diverse male Veteran population.

*Performance in Veterans at Higher Risk*

HIV infection: In the study by Yin et al[26] that examined a modified FRAX score in male Veterans with and without HIV infection (n = 24,451), FRAX (not including BMD, and assuming low risk of parental hip fracture) underestimated true fracture risk with a goodness-of-fit observed/expected (O/E) ratio for MOF of 1.39 (95% CI 1.30 to 1.48) in the total cohort, and lower accuracy for those with HIV infection (O/E of 1.62, 95% CI 1.45 to 1.81). While inclusion of HIV as a cause of secondary osteoporosis ("yes" in FRAX) improved its accuracy in predicting MOF (O/E 1.20) in HIV infection, it still underestimated risk in this population. As seen in Table 3, even when HIV was included as a cause of secondary osteoporosis, FRAX was a poor predictor of hip fracture among male Veterans with (O/E 2.66) and without (O/E 3.56) HIV infection. Overall, these data suggest modified FRAX (without BMD) underestimates fracture risk in HIV. It is worth noting that sensitivity for detecting MOF and hip fractures were very low in this study, both in HIV infected and uninfected populations. Specifically, sensitivity for detecting fracture was only 0 to 6.4% using accepted FRAX thresholds for initiating osteoporosis care. It is possible that younger age of Veterans (mean = 55.6 years) contributed to this finding, given mean age was substantially different than other studies we identified (although FRAX is validated in ages 40–60), or that EHR-based ascertainment of certain FRAX components (*eg*, smoking, alcohol) limited its performance in this population. Also, this study included a more racially diverse population than Williams et al, with 46.3% and 8.7% Veterans identifying as Black race or Hispanic ethnicity, respectively. There may be differences in the ability of FRAX to predict fracture across races and ethnicities; however, current data are limited in this regard and no conclusions can be drawn.

HCV infection and HCV/HIV co-infection: Yin et al[26] also examined the performance of modified FRAX (without BMD) in predicting fracture among male Veterans with HCV infection, regardless of HIV status. As detailed in Table 3, when patients with HCV infection—without HIV—were assumed to have secondary osteoporosis in the modified FRAX, the accuracy of the tool for predicting MOF (O/E 1.27) was similar to that seen in the HIV

population with or without HCV infection (O/E 1.20). The risk of MOF was also underestimated by modified FRAX in the HCV/HIV co-infected population (O/E 1.48); it was likewise a poor predictor of hip fracture among male Veterans with HCV infection, independent of HIV status (O/E for HIV+ 3.87, O/E for HIV- 3.44).

Androgen deprivation therapy (ADT): Another population of special interest examined by FRAX was male Veterans who had undergone ADT, a known risk factor for accelerated bone loss.[48] Of 115 patients in this study (60% African American), 54% met standard thresholds for FRAX (10-year risk of ≥ 20% for MOF or ≥ 3% of hip fracture), and only 33% had a T-score in the osteoporosis range via DXA. The proportion of Veterans with osteoporosis who were captured by these FRAX thresholds is unclear based on available data. Authors calculated a "FRAX" T-score by using the FRAX patch, which was available through the National Osteoporosis Foundation (NOF) website. Comparison of the "FRAX" T-score (mean –0.4 (±1.5) to BMD at the forearm (–1.2, ±1.9) and hip (–1.4, ±1.1) suggests that FRAX underestimates the risk of osteopenia and/or osteoporosis at these sites in men treated by ADT. In contrast, the "FRAX" T-score approximated BMD at the spine (0.0, ±1.8), although this site is known to be less predictive of fracture in men as compared to women.[81-83] While far from definitive, these data suggest that FRAX underestimates osteoporosis at the wrist and hip in men treated with ADT.

## *OST*

### *Performance in Average Risk Veterans*

Three studies[28,53,54] (n = 1,162) examined OST with the intent of evaluating its performance in a community-dwelling male Veteran population. OST predicted osteoporosis similarly among general populations (AUC 0.632 to 0.740) as among Veterans (AUC 0.670 to 0.890). Williams et al[28] found an OST score below 1 to predict osteoporosis by DXA with a sensitivity of 69% and specificity of 60% (AUC 0.71, 95% CI 0.65 to 0.76). The same cut-off of 1 in Adler et al[54] had higher sensitivity (75%) and specificity (80%) for identifying osteoporosis (AUC 0.84, 95% CI 0.75 to 0.92), and sensitivity increased substantially (93%) when an OST cut-off of 3.0 was used. Last, Richards et al found an OST score below 6 to have good sensitivity (82.6%), but a much lower specificity (33.6%) (AUC 0.67, 95% CI NR).[53] There are important differences between these 3 study populations that may account for variation in the predictive ability of OST. Williams et al (n = 463) and Richards et al (n = 518) enrolled Veterans from the general medicine setting (and had similar AUC of 0.67 to 0.71), and Adler et al recruited a smaller cohort (n = 181) of Veterans from pulmonary and rheumatology clinics, where the risk of osteoporosis was likely higher. While Adler et al[54] found that OST performed similarly in patients with (AUC 0.79, 95% CI NR) and without (AUC 0.80, 95% CI NR) current glucocorticoid use, they did not examine the role of prior glucocorticoid exposure or the presence of inflammatory conditions that were likely driving heightened osteoporosis risk in this population. One study[53] noted OST to have a better predictive ability in non-Hispanic Caucasian (AUC 0.72, 95% CI NR) versus African American (AUC 0.58, 95% CI NR) Veterans, and in older age groups (*ie*, > 65 years of age). While another study noted especially high AUC (0.99) in Veterans aged ≥ 80 years, there was no consistent trend observed in OST performance across age subgroups.[54] Likewise, no notable differences across racial/ethnic subgroups were noted in 2 of 3 studies,[28,54] even though only 1 had a sufficient African American population (25.1% of total cohort) to examine this association (*ie*, Williams et al study was 94.2% Caucasian). With the exception of 1 study where the reference population was not reported,[28] all studies examining

OST utilized the male reference population from NHANES III to define osteoporosis at the hip. See Table 3 for additional detail on these studies.

A separate study focused on the performance of OST in an African American population. Sinnott et al[51] examined 128 African American male Veterans recruited from general medicine clinics and found OST to perform reasonably well at a cut-off of 4, with a sensitivity of 89% and specificity of 54% for detecting osteoporosis (AUC 0.89, 95% CI 0.75 to 1.03). Although African American males have traditionally been considered a low-risk group for osteoporosis, a prevalence of 7% was noted in this study; this is unexpectedly high, but it may still be an underestimation, given a Caucasian male normative database (NHANES III) was used for the hip, and Caucasian men are known to have lower BMD than African American men.

*Performance in Veterans at Higher Risk*

One study[56] examined the performance of OST in 282 male Veterans with rheumatoid arthritis, and a score below 4 predicted osteoporosis with a sensitivity of 78% and specificity of 45% (AUC not reported).

**Table 3. FRAX, Modified FRAX and OST for Assessing Risk of Osteoporosis and/or Fracture in Male Veterans**

| Study<br>Risk assessment tool<br>(Tool components) | N with outcome<br>Total n<br>(Condition-specific population) | Outcome metric (definition)<br>Reference population | Sens & Spec<br>(Threshold) | AUC (95% CI) | ROB |
|---|---|---|---|---|---|
| Williams, 2017[28]<br>e-FRAX<br>FRAX adapted to EHR (age, sex, weight, height, previous fracture, parental hip fracture, smoking, glucocorticoid treatment, rheumatoid arthritis, alcohol intake) | Osteoporosis: 112<br>Total n = 463 | Osteoporosis (T-score ≤ –2.5)<br>NR | Sens 0.688<br>Spec 0.544<br>(≥ 20 % risk for MOF; ≥ 3 % risk for hip fracture) | AUC/ROC 0.65 (0.59 to 0.71) | Low risk of bias |
| Williams, 2017[28]<br>FRAX<br>(age, sex, weight, height, previous fracture,parental hip fracture, smoking, glucocorticoid treatment, rheumatoid arthritis, alcohol) | | Osteoporosis (T-score ≤ –2.5)<br>NR | Sens NR<br>Spec NR<br>(NOF and ACR) FRAX : ≥ 3% for hip fracture; ≥ 6.5% for major osteoporotic fractures | AUC/ROC 0.72 (0.67 to 0.78) | |
| Yin, 2016[26]<br>Modified FRAX (total, HIV+, HIV-) | MOF HIV-: 609<br>MOF HIV+: 326<br>Hip HIV-: 148<br>Hip HIV+: 93<br>Total n = 24451 | Major osteoporotic fracture rate defined by ICD9 codes | NR<br>NR | Observed/expected: total 1.39 (95% CI 1.30 to 1.48); HIV- 1.29 (95% CI 1.19 to 1.40); HIV+ 1.62 (95% CI 1.45 to 1.81); p-value for difference between HIV+ vs HIV- O/E 0.03 | At risk of bias |

| Study<br>Risk assessment tool<br>(Tool components) | N with outcome<br>Total n<br>(Condition-specific population) | Outcome metric (definition)<br>Reference population | Sens & Spec<br>(Threshold) | AUC (95% CI) | ROB |
|---|---|---|---|---|---|
| Yin, 2016[26]<br>Modified-FRAX calculated with HIV as secondary osteoporosis | (HIV+ n = 7064<br>HIV- n = 17387) | Major osteoporotic fracture rate defined by ICD9 codes | NR<br>NR | Observed/expected: HIV- 1.29 (95% CI 1.19 to 1.40); HIV+ 1.20 (95% CI 1.08 to 1.34) | |
| Yin, 2016[26]<br>Modified-FRAX (total, HIV infected, HIV-) | | Hip fracture rate defined by ICD9 codes | NR<br>NR | Observed/expected hip fracture: total 3.87 (95% CI 3.42 to 4.40); HIV- 3.56 (95% CI 3.03 to 4.18); HIV+ 4.52 (95% CI 3.68 to 5.53) | |
| Yin, 2016[26]<br>Modified-FRAX calculated with HIV as secondary osteoporosis (HIV+, HIV-)[b] | | Hip fracture rate defined by ICD9 codes | NR<br>NR | Observed/expected: HIV- 3.56 (95% CI 3.03 to 4.18); HIV+ 2.66 (95% CI 2.17 to 3.26) | |
| Yin, 2016[26]<br>Modified-FRAX calculated with HIV as secondary osteoporosis: HIV+ | | Major osteoporotic fracture rates defined by ICD9 codes | Sens 6.4%<br>Spec 98.6%<br>(Since none met the NOF fracture threshold of > 20%, the age-specific thresholds endorsed by European osteoporosis societies was utilized [6.3% to 13.4% in 50- to 70-year-olds]) | NR | |
| Yin, 2016[26] | | Major osteoporotic fracture rates defined by ICD9 codes | Sens 2.6%<br>Spec 99.5% | NR | |

| Study<br>Risk assessment tool<br>(Tool components) | N with outcome<br>Total n<br>(Condition-specific population) | Outcome metric (definition)<br>Reference population | Sens & Spec<br>(Threshold) | AUC (95% CI) | ROB |
|---|---|---|---|---|---|
| Modified-FRAX calculated with HIV as secondary osteoporosis: HIV- | | | (Since none met the NOF fracture threshold of > 20%, the age-specific thresholds endorsed by European osteoporosis societies was utilized [6.3% to 13.4% in 50- to 70-year-olds]) | | |
| Yin, 2016[26]<br>Modified-FRAX calculated with HIV as secondary osteoporosis: HIV+ | | Hip fracture rate defined by ICD9 codes | Sens 3.2%<br>Spec 99.0%<br>(FRAX: ≥3 % for hip fracture probability) | NR | |
| Yin, 2016[26]<br>Modified-FRAX calculated with HIV as secondary osteoporosis: HIV- | | Hip fracture rate defined by ICD9 codes | Sens 0%<br>Spec 99.9%<br>(FRAX: ≥3 % for hip fracture probability) | NR | |
| Yin, 2016[26]<br>Modified-FRAX calculated with HCV as cause of secondary osteoporosis: HIV+ | | Major osteoporotic fracture rate defined by ICD9 | NR | Observed/expected: 1.48 (1.33 to 1.65) | |

| Study<br>Risk assessment tool<br>(Tool components) | N with outcome<br>Total n<br>(Condition-specific population) | Outcome metric (definition)<br>Reference population | Sens & Spec<br>(Threshold) | AUC (95% CI) | ROB |
|---|---|---|---|---|---|
| Yin, 2016[26]<br>Modified-FRAX calculated with HCV as cause of secondary osteoporosis: HIV- | | Major osteoporotic fracture rate defined by ICD9 | NR | Observed/expected: 1.27 (1.17 to 1.37) | |
| Yin, 2016[26]<br>Modified-FRAX calculated with HCV as cause of secondary osteoporosis: HIV+ | | Hip fracture rate defined by ICD9 | NR | Observed/expected: 3.87 (3.16 to 4.75) | |
| Yin, 2016[26]<br>Modified-FRAX calculated with HCV as cause of secondary osteoporosis: HIV- | | Hip fracture rate defined by ICD9 | NR | Observed/expected: 3.44 (2.93 to 4.04) | |
| Adler, 2010[48]<br>FRAX (age, sex, weight, height, previous fracture, parental hip fracture, smoking, glucocorticoid treatment, rheumatoid arthritis, alcohol) | Osteoporosis: 33%<br>Total n = 115<br>(Androgen deprivation therapy) | Osteoporosis (T-score ≤ -2.5)<br>NHANES III (male) | Sens NR<br>Spec NR<br>FRAX: ≥ 20 % risk for major osteoporotic fracture or ≥ 3% risk for hip fracture | 54% had a FRAX probability above 20% for MOF or 3% for hip fracture[b]<br>35% had a T-score ≤ -2.5 measured with DXA[b] | Low risk of bias |
| Williams, 2017[28] | Osteoporosis: 112 | Osteoporosis (T-score ≤ -2.5) | Sens .688<br>Spec .598 | AUC/ROC 0.71 (0.65 to 0.76) | Low risk of bias |

| Study<br>Risk assessment tool<br>(Tool components) | N with outcome<br>Total n<br>(Condition-specific population) | Outcome metric (definition)<br>Reference population | Sens & Spec<br>(Threshold) | AUC (95% CI) | ROB |
|---|---|---|---|---|---|
| OST<br>(Age, Weight) | Total n = 463 | NR | (OST score <0.99) | | |
| Richards, 2014[53]<br>OST<br>(Age, Weight) | Osteoporosis: 92<br>Total n = 518 | Osteoporosis (T-score ≤ -2.5)<br>NHANES III (male) | Sens 82.6%<br>Spec 33.6%<br>(OST score ≤ 6) | AUC/ROC 0.67 | At risk of bias |
| Sinnott, 2006[51]<br>OST<br>(Age, Weight) | Osteoporosis: 7%<br>N = 128 | Osteoporosis (T-score ≤ -2.5)<br>Caucasian male normative database for the hip and the manufacturer's female spine database | Sens 89<br>Spec 54<br>(OST score 4) | AUC/ROC 0.89 (95% CI 0.75 to 1.03) | At risk of bias |
| Richards, 2009[56]<br>OST<br>(Age, Weight) | Osteoporosis: 50<br>Total n = 282 (Rheumatoid arthritis) | Osteoporosis (T-score ≤ -2.5)<br>NHANES III (male) | Sens 78<br>Spec 45<br>(OST score ≥ 4) | NR | At risk of bias |
| Richards, 2009[56]<br>OST<br>(Age, Weight) | | Osteoporosis (T-score ≤ -2.5)<br>NHANES III (male) | Sens 6<br>Spec 94<br>(OST ≤ -2) | NR | |
| Richards, 2009[56] | | Osteopenia (T-score between -1.0 and -2.5) | Sens 64<br>Spec 54 | NR | |

| Study<br>Risk assessment tool<br>(Tool components) | N with outcome<br>Total n<br>(Condition-specific population) | Outcome metric (definition)<br>Reference population | Sens & Spec<br>(Threshold) | AUC (95% CI) | ROB |
|---|---|---|---|---|---|
| OST<br>(Age, Weight) | | NHANES III (male) | (OST ≤ 4) | | |
| Adler, 2003[54]<br>OST<br>(Age, Weight) | Osteoporosis: 15.6%<br>Total n= 181<br>(Pulmonary or rheumatology clinic population) | Osteoporosis (T-score ≤ -2.5)<br>NHANES data for hip, Hologic reference source for spine | Sens 93<br>Spec 66<br>(OST score 3) | AUC/ROC 0.836 | At risk of bias |
| Adler, 2003[54]<br>OST<br>(Age, Weight) | | Osteoporosis (T-score ≤ -2.5)<br>NHANES data for hip, Hologic reference source for spine | Sens 82<br>Spec 74<br>(OST score 2) | | |
| Adler, 2003[54]<br>OST<br>(Age, Weight) | | Osteoporosis (T-score ≤ -2.5)<br>NHANES data for hip, Hologic reference source for spine | Sens 75<br>Spec 80<br>(OST score 1) | | |

[a]HIV+ are people living with HIV and HIV- are people not identified as living with HIV.
[b]Exact percentages obtained via correspondence with the author.

### *Mscore*

Mscore was assessed in 1 study[67] of male Veterans who were recruited primarily from general medicine clinics, and to a lesser extent, endocrinology and osteoporosis clinics. They found an Mscore cut-off of 9 predicted osteoporosis (T-score on DXA ≤ -2.5) with 88% sensitivity and 57% specificity (AUC 0.84, 95% CI 0.74 to 0.95) in the validation cohort of Caucasian male Veterans (n = 197). The same cut-off in a reduced version of Mscore including only age and weight yielded similar sensitivity and specificity, 85% and 58% respectively (AUC 0.81, 95% CI 0.69 to 0.92). The reduced Mscore performed even better in a validation cohort of African American Veterans (n = 134), with a sensitivity as high as 93%, a specificity of 79%, and an AUC of 0.89 (95% CI 0.79 to 0.98) when race-specific reference data (NHANES III) were used (AUC 0.99, 95% CI 0.98 to 1.01, if Caucasian reference data used). Combined with the robust performance of OST in this study, these data suggest age and weight may be the most influential clinical variables in risk assessment tools for osteoporosis, and that the 5-variable Mscore may be more complex than necessary to assess risk of osteoporosis in male Veterans.

### *VA-FARA*

One study[28] examined the use of VA-FARA in 463 male Veterans. Similar to FRAX, VA-FARA calculates 10-year probabilities of MOF and hip fracture. Using the same cut-offs as FRAX for these outcomes (see Table 4), Williams et al found it had a sensitivity of 64.3% and specificity of 58.4% for predicting osteoporosis (AUC 0.64, 95% CI 0.58 to 0.70). It suffers from similar drawbacks as with e-FRAX noted above (whose AUC was similar at 0.65 [95% CI 0.59 to 0.71]), as it relies on EHR data that may not be accurate or complete. Furthermore, it does not appear to outperform simpler tools, such as OST or the reduced Mscore.

Both Mscore and VA-FARA are "homegrown" tools examined in 1 study each[28,67] (as with e-FRAX described above), so while Mscore appears to better predict osteoporosis than VA-FARA, there are insufficient data to recommend 1 of these approaches over another or to recommend any of these tools above FRAX and OST that have been studied and validated across broader populations.

**Table 4. Studies Examining Mscore and VA-FARA as One-off Tools to Predict Osteoporosis and/or Fracture in Male Veterans**

| Study<br><br>Risk Assessment tool<br><br>(Tool components) | N with outcome<br><br>Total n<br><br>(Condition-specific population) | Outcome metric (definition)<br><br>Reference population | Sens & Spec<br><br>(Threshold) | AUC/ROC (95% CI) | ROB |
|---|---|---|---|---|---|
| Zimering, 2007[67]<br><br>Mscore | Osteoporosis in validation cohort: 11%<br><br>Validation group 1 n = 197 | Osteoporosis (T-score ≤ -2.5)<br><br>NHANES III Male | Sens 88<br>Spec 57<br><br>(MSCORE of 9) | AUC/ROC 0.84 (NR) | At risk of bias |

| Study<br><br>Risk Assessment tool<br><br>(Tool components) | N with outcome<br><br>Total n<br><br>(Condition-specific population) | Outcome metric (definition)<br><br>Reference population | Sens & Spec<br><br>(Threshold) | AUC/ROC (95% CI) | ROB |
|---|---|---|---|---|---|
| (Age, weight, gastrectomy, emphysema, 2 or more prior fractures) | | | | | |
| Zimering, 2007[67]<br><br>Mscore<br><br>(Age, weight) | Osteoporosis in African American validation group: 11%<br><br>African American validation group n = 134 | Osteoporosis (T-score ≤ -2.5)<br><br>NHANES III Male (race-specific) | Sens 93<br>Spec 79<br><br>(Mscore of 9) | AUC/ROC 0.89 (95% CI 0.79 to 0.98) | At risk of bias |
| Zimering, 2007[67]<br><br>Mscore<br><br>(Age, weight) | Osteoporosis in Caucasian validation cohort: 11%<br><br>Validation group 1 n = 197 | Osteoporosis (T-score ≤ -2.5)<br><br>NHANES III Male (race-specific) | Sens 85<br>Spec 58<br><br>(Mscore of 9) | AUC/ROC 0.81 (95% CI 0.69 to 0.92) | At risk of bias |
| Williams, 2017[28]<br><br>VA-FARA<br><br>(Prior fracture, age > 80, weight, DM complications, malnutrition, CVA, smoking, EtOH, 6–12 clinic visits, 13+ clinic visits in prior year, fall risk) | Osteoporosis: 112<br><br>Total n = 463 | Osteoporosis (T-score ≤ -2.5)<br><br>NR | Sens .643<br>Spec .584<br><br>(≥ 3% for hip fracture; ≥ 20% for major osteoporotic fractures) | AUC/ROC 0.64 (0.58 to 0.70) | Low risk of bias |

## Assessment of Individual Risk Factors

Of the 12 studies[51,69-79] that examined individual risk factors for low BMD and/or fracture among male Veterans (n = 585,400), 9 studies[51,69,71,73,75-79] exclusively examined the role of medical conditions as risk factors (*ie,* HIV infection, osteomyelitis, elevated BMI, chronic kidney disease, vitamin D deficiency, chronic pancreatitis), while the remaining studies assessed exposures (*eg,* medication use), or combinations of conditions and exposures in specific Veteran

populations. These are summarized in Table 5. Overall, there was a high level of heterogeneity in terms of: a) how risk factors were defined; b) the level of detail regarding these risk factors; c) populations in which risk factors were studied; d) outcomes; and e) how effect sizes were calculated and reported across studies (*eg,* odds ratios, hazard ratios, correlation coefficients, F-statistic).

### *Studies of Medical Conditions as Risk Factors*

*Osteomyelitis, HIV infection and HIV-associated factors in the VA Aging Cohort Study (VACS)*

Four studies examined risk factors for fracture using data from the VACS that includes Veterans with HIV infection (+/- HCV infection), as well as age- and race-matched Veteran controls without HIV infection.[71,75,77,79]

One study[71] investigated the risk of fragility fracture in male Veterans with and without osteomyelitis, and after adjusting for demographic and clinical variables (including steroid use, 9 of the 11 FRAX variables, as well as HIV, HCV, and diabetes status), they found presence of osteomyelitis to be associated with highest odds of vertebral fracture (aOR 2.43, 95% CI 1.17, 5.03), followed by upper arm/humerus fracture (aOR 1.95, 95% CI 1.02, 3.74) and fracture at any site (aOR 1.65, 95% CI 1.15, 2.36); no significant association was noted with hip fracture. A sensitivity analysis excluding patients with pelvic or lower extremity osteomyelitis was conducted as a means of accounting for fall risk related to osteomyelitis at these sites and similar results were noted. Another potential source of confounding is the exclusion of nearly half of male Veterans aged 50–70 from this database due to incomplete data for fracture-associated variables; the eligible cohort decreased to 24,251 from 42,924 due to incomplete data. Lack of understanding of how this population differed from the analytic cohort in terms of osteomyelitis and fracture risk limits conclusions that can be drawn.

One study[77] utilizing VACS data examined HIV infection as a risk factor for fragility fracture (composite of hip, spine, humerus) in male Veterans (n = 40,115). While they found HIV to predict fracture, (aHR 1.24, 95% CI 1.11 to 1.39), this was not significant after adjusting for BMI in multivariable analysis. This may suggest that maintenance of weight and nutritional status is protective against fracture among male Veterans with HIV infection. To explore how HIV severity and care factors may impact fracture risk, another study[75] in this cohort examined how a "VACS index" score and its individual components related to fracture risk in HIV infection. The VACS index includes demographic, clinical, and laboratory data, as well as HIV-specific variables such as CD4 count, HIV viral load, and use of antiretroviral medications relevant to bone health, specifically tenofovir, protease inhibitors, and efavirenz. A higher VACS index approximates frailty in male Veterans with HIV infection and predicts mortality.[84-86] Notably, this study found only a modest association between the VACS index and fragility fracture (aHR 1.15, 95% CI 1.11 to 1.19) – this was also true of the individual components of the index, including use of antiretroviral therapies that may have direct effects on bone quality (tenofovir, protease inhibitors) or raise the risk of falls through central nervous system effects (efavirenz). Increasing age appeared to be the most potent risk factor for fragility fracture in this population (aHR for age by 10-year increments, 1.40, 95% CI 1.27 to 1.54), not unlike populations without HIV infection.

Another study evaluated the VACS Index and risk of low BMD among HIV-positive Veterans.[79] The VACS Index score was found to be significantly associated with the risk of low BMD (*ie*, osteopenia or osteoporosis by DXA) in HIV-positive Veterans, with the odds of low BMD increasing 1.21 times for every 10 unit increase in VACS index score. However, as per Womack et al (a much larger cohort),[77] a higher VACS score may not necessarily translate to increased fracture risk.

*Weight (kg) and body mass index (BMI)*

In a cohort of 128 African American male Veterans with average BMI in the overweight range (28.9 kg/m$^2$), weight below 85kg was found to predict low BMD with a sensitivity of 74% and specificity of 50%.[51] Specifically, weight (< 85kg) predicted osteoporosis (T-score ≤ -2.5) with an AUC of 0.75 (95% CI 0.57 to 0.92), and unexpectedly, BMI was not as predictive (AUC 0.67, 95% CI 0.47 to 0.87).

*Chronic kidney disease (CKD)*

One study examined the association between different CKD stages and the risk of fracture.[73] Of the 712,918 male Veterans in this study, the vast majority (95.2%) had CKD stage 3. CKD stages 4 and 5 accounted for 4.3% and 0.5% of the cohort, respectively. The outcome of fracture included any site, and 22.6% of all fractures in this study occurred in the rib or clavicle, which are not sites of fragility that define clinical osteoporosis. However, most other fractures occurred in either the hip/femur (25.8%) or vertebra (16.9%). A critical finding of this study is the importance of accounting for mortality in CKD populations, as the authors noted that the association between CKD and fracture was no longer significant after accounting for death as a competing event in this male Veteran population. The only exception was for CKD stage 3, where there was a modestly elevated OR of 1.07 (95% CI 1.02 to 1.11) for predicting fracture, even after adjusting for mortality. Age and race did not moderate the effects of CKD on fracture in this study.

*Vitamin D deficiency*

Vitamin D[78] is a critical contributor to bone metabolism and its deficiency, as defined by a 25-hydroxy vitamin D level < 15 ng/mL, was explored as a risk factor for low BMD in 1 study of 112 African American male Veterans. Prevalence of osteoporosis by T-scores (as per WHO criteria) was similar in Veterans with (3.5%, n = 58) and without (3.7%, n = 54) vitamin D deficiency. In contrast, osteopenia was more prevalent in patients without vitamin D deficiency (25.9%) as compared to those with vitamin D deficiency (19%). In this study, the authors noted no correlation between vitamin D level and BMD (g/cm$^2$).

*Chronic pancreatitis*

Chronic pancreatitis is relevant to bone health because exocrine deficiency from pancreatitis may lead to malabsorption of nutrients (including vitamin D) and a heightened risk of fracture.[69] In a retrospective analysis of 3,257 male Veterans diagnosed with chronic pancreatitis, 4.7% (n = 153) sustained a fracture at any site over a 10-year period. Notably, Veterans were identified as having chronic pancreatitis by diagnosis code, which has been shown to have a sensitivity of 87% and specificity of 86% in VA databases.[87] When compared to an unmatched control population of similar age (~54 years) but without chronic pancreatitis, the presence of chronic pancreatitis was associated with significantly higher odds of hip fracture (aOR 2.69, 95% CI 2.13

to 3.40), followed by fracture at any site (OR 1.73, 95% CI 1.46 to 2.05) and vertebral fracture (aOR 1.56, 95% CI 1.06 to 2.31). No significant impact was noted on the risk of wrist fracture. Authors adjusted for age and etiology of pancreatitis (alcohol, smoking, or both) in multivariable analysis, but did account for race/ethnicity, medication use (*eg,* use of proton pump inhibitors, steroids, bisphosphonates), severity of pancreatic disease, or aspects of chronic pancreatitis treatment.

*Post-traumatic stress disorder (PTSD)*

Veterans who have been prisoners of war (POW) may have multiple risk factors for bone loss related to captivity, including dietary deficiencies, vitamin D deficiency, and immobility. Evidence suggests that repatriation of POW is associated with reversal of bone loss in most cases[88]; however, the burden of PTSD remains high in this population. One study[76] evaluated the effect of PTSD by comparing BMD of repatriated POW with (n = 61) and without (n = 180) PTSD, and a control group (n = 79) of combat-experienced, non-POW Veterans without PTSD. The mean age was 62.2–63.4 years in this study, and mean captivity duration for repatriated POW (regardless of PTSD status) was 53 months. Estimated weight loss was similar between POW, regardless of PTSD status (PTSD+ 44.5 lbs; PTSD- 43.6 lbs). Despite noting a significant difference across groups in terms of hip T-scores (lower in POW PTSD+ vs POW PTSD- and control group), a similar pattern was not noted for spine T-scores, and all Veterans had T-scores in the "normal" range (*ie,* > -1.5). Furthermore, while age, BMI, ethnicity, and alcohol use were adjusted for in the model, smoking status was not accounted for. It is plausible that smoking elevated the risk for bone loss in POW with PTSD as their tobacco exposure by pack-years was notably higher than the other 2 groups (21 pack-years vs 16 to 17 pack-years).

### *Exposures or Combinations of Medical Conditions and Exposures as Risk Factors*

*Thyroid cancer and levothyroxine supplementation*

In the setting of thyroid cancer, high doses of thyroid hormone (levothyroxine) are purposefully given to suppress TSH and limit growth of cancer cells.[70] High levels of thyroid hormone increase bone turnover and promote bone loss, and in non-Veteran populations, an association has been noted between treatment of thyroid cancer (with high levothyroxine doses) and osteoporosis.[89,90] One case-control study examined this association among male Veterans by comparing osteoporosis and fracture rates (defined by diagnosis codes) in patients with and without thyroid cancer, matched for age, sex, weight, and steroid use. Both cases (thyroid cancer) and controls (hypothyroidism) were on levothyroxine therapy, however, Veterans with thyroid cancer were being treated to lower TSH goals as part of their management of thyroid cancer, whereas controls were treated to euthyroidism. The odds of being assigned an osteoporosis diagnosis were higher among Veterans with thyroid cancer versus controls (OR 1.46, 95% CI 1.26 to 1.68). For unclear reasons, diagnosis of fracture was significantly lower in thyroid cancer cases versus controls (OR 0.70, 95% CI 0.58 to 0.85). The authors did not observe a difference between groups in receipt of osteoporosis medications in their multivariable model. These data are limited by uncontrolled confounders (*eg,* smoking, alcohol), and since the prescription of high-dose levothyroxine was likely to have prompted providers to screen for osteoporosis, ascertainment bias is highly likely to have influenced these results.

*Ulcerative colitis and risks related to malnutrition, vitamin D deficiency, steroid use, and others*

Several studies in non-Veteran populations have identified inflammatory bowel disease (including ulcerative colitis) as a risk factor for low BMD and fractures.[91] One nationwide VA study assessed individual risk factors for low BMD and fracture among 34,665 Veterans with ulcerative colitis; diagnosis codes were used to define outcomes of osteopenia and osteoporosis (collectively defined as "low BMD" in this study) and fragility fracture.[74] Risk factors were also defined by diagnosis codes, except for prednisone use that required pharmacy data. The presence of tobacco abuse, malnutrition, vitamin D deficiency, and high prednisone exposure (> 11,136 mg over a 10-year period) were associated with higher risk of both low BMD and fragility fracture among male Veterans. High prednisone exposure was strongly linked to low BMD (OR 8.9, 95% CI, 7.8 to 10.2) and appeared to follow a predictable dose-response trend, whereas it was associated with fracture to a lesser extent (OR 1.8, 95% CI 1.3 to 2.5) and ORs did not consistently increase with greater prednisone exposure. It is also important to note that this study included osteopenia in the outcome of low BMD, which may have weakened the relationship with examined risk factors.

### *Antipsychotic Drugs*

Use of antipsychotic medications may lead to hyperprolactinemia, which can cause hypogonadism, a known risk factor for bone loss. They may also contribute to orthostatic hypotension and fall risk. We identified 1 study that examined the association between antipsychotic use (≥ 3 months over 10-year period) and fracture rates in a predominantly male Veteran population (91% male, n = 5,824).[72] This was done by comparing fracture rates (not limited to fragility fracture) between Veterans with antipsychotic use and a control group of Veterans (not on antipsychotics) matched for several key comorbidities, including: heart failure, emphysema, depression, diabetes, end-stage renal disease, and schizophrenia. Smoking and alcohol use were not adjusted for in multivariable analysis, and these exposures were significantly more prevalent in Veterans on antipsychotic medications. Despite this, prevalence of fracture was similar between Veterans with (10.0%) and without (10.2%) long-term antipsychotic use in this study, and authors found no significant association with fracture.

**Table 5. Assessment of Individual Risk Factors**

| Author, Year | Risk factors<br>Total n | Prevalence OR | Outcome metric | Effect size (95% CI) | ROB |
|---|---|---|---|---|---|
| Munigala, 2016[69] | Chronic pancreatitis<br>Total n = 3079 | 153 fractures/ 3257 people with chronic pancreatitis | Fracture at all sites | adjusted OR 1.73% (1.46 to 2.05)<br>p < 0.0001 | High risk |
| | | | Hip fracture | adjusted OR 2.69% (2.13 to 3.40)<br>p < 0.0001 | |
| | | | Vertebral fracture | adjusted OR 1.56% (1.06 to 2.31)<br>p = 0.0257 | |

| Author, Year | Risk factors<br>Total n | Prevalence OR | Outcome metric | Effect size (95% CI) | ROB |
|---|---|---|---|---|---|
| | | | Wrist fracture | adjusted OR 1.18% (0.90 to 1.55)<br>p = 0.2382 | |
| | Chronic pancreatitis (age 45–65)<br>n = 2038 | | Fracture at all sites | OR 2.41 (1.96 to 2.96)<br>p < 0.0001 | |
| | Chronic pancreatitis (age > 65)<br>n = 525 | | Fracture at all sites | OR 2.75 (1.99 to 3.80)<br>p < 0.0001 | |
| Hsieh, 2019[71]<br>VACS cohort | Osteomyelitis<br>n = 24451 | 6.5% fracture in men with osteomyelitis; 3.8% in men without osteomyelitis | All fracture (hip, upper arm, vertebra) | Adjusted OR 1.649 (1.154 to 2.356) | Unclear |
| | Osteomyelitis<br>n = 24451 | | Hip fracture | Adjusted OR 1.762 (0.944 to 3.289) p 0.08 | |
| | Osteomyelitis<br>n = 24451 | | Upper arm fracture | Adjusted OR 1.95 (1.016 to 3.744) p 0.04 | |
| | Osteomyelitis<br>n = 24451 | | Vertebral fracture | Adjusted OR 2.428 (1.173 to 5.029), p 0.02 | |
| Hall, 2018[73] | CKD stage 3<br>n = 339,278 | 12.4% without CKD had fractures over median 5.2 years follow up vs 15.7% of those with CKD | Fracture by ICD and CPT codes | Cox proportional adjusted HR 0.95 (0.91 to 0.99)<br>Fine and Gray[a] Adjusted subdistribution HR 1.07 (1.02 to 1.11) | Unclear |
| | CKD stage 4<br>n = 15,167 | | Fracture by ICD and CPT codes | Cox proportional adjusted HR 1.32 (1.16 to 1.49)<br>Fine and Gray[a] Adjusted subdistribution HR 1.07 (0.94 to 1.22) | |
| | CKD stage 5<br>n = 2,014 | | Fracture by ICD and CPT codes | Cox proportional adjusted HR 1.91 (1.45 to 2.50) | |

| Author, Year | Risk factors<br>Total n | Prevalence OR | Outcome metric | Effect size (95% CI) | ROB |
|---|---|---|---|---|---|
| | | | | Fine and Gray[a] Adjusted subdistribution HR 1.31 (0.97 to 1.77) | |
| Shahani, 2019[79]<br>VACS cohort | VACS frailty index | 109 (56%) had osteoporosis or osteopenia out of 195 | DXA T-score or Z-score<br>Reference population NR | Odds of low BMD increasing 1.21 times for each 10-unit increase in VACS Index score [confidence interval (95% CI) 1.03–1.42; p = .02] | High risk |
| Womack, 2013[75]<br>VACS cohort | VACS frailty index<br>n = 40,115 | 588 first fragility fractures (210 hip, 111 vertebral, 267 upper arm) | Fragility Fracture (composite of hip, vertebral and upper arm fractures, defined using ICD-9 codes) | HR for covariate-adjusted association between fragility fracture and VACS Index Score (for 10-unit increase in index): 1.15 (1.11 to 1.19) | Unclear |
| Hain, 2011[76] | PTSD<br>320 (only 61 had PTSD, 19.1%) | NR | Total hip T-score<br>Reference population NR | F (2,313) = 3.02, p < 0.05, partial $h^2$ = 0.02 | High risk |
| | PTSD<br>n = 320 (only 61 had PTSD, 19.1%) | | Total spine T-scores<br>Reference population NR | F (2,313) = 1.54, p < 0.22, partial $h^2$ = 0.01 | |
| Womack, 2011[77]<br>VACS cohort | HIV<br>n = 119,318 men, 33% of whom were HIV infected | 1615 first fractures (496 hip, 322 vertebral, and 797 upper arm fractures) | ICD9 based fragility fracture of hip, vertebrae, or upper arm | Adjusted HR (without BMI): 1.24 (1.11, to 1.39)<br>Adjusted HR (including BMI and BMI squared): 1.10 (0.97 to 1.25) | Unclear |
| Akhter, 2009[78] | Vitamin D 25-OHD > 15 ng/ml (Group I)<br>n = 54 | Group I: 3.7% with osteoporosis | BMD of spine | Correlation coefficient r = 0.13, p >0.05 | Unclear |

| Author, Year | Risk factors<br>Total n | Prevalence OR | Outcome metric | Effect size (95% CI) | ROB |
|---|---|---|---|---|---|
| | Vitamin D 25-OHD > 15 ng/ml (Group I)<br>n = 54 | Group II: 3.5% with osteoporosis<br>Group I: 25.9% with osteopenia<br>Group II: 19% with osteopenia | BMD of hip | Correlation coefficient, r = 0.18, p>0.05 | |
| | Vitamin D 25-OHD ≤ 15 ng/ml (Group II)<br>n = 58 | | BMD of spine | Correlation coefficient r = 0.26, p = 0.05 | |
| | Vitamin D 25-OHD ≤ 15 ng/ml (Group II)<br>n = 58 | | BMD of hip | Correlation coefficient: r = 0.27, p < 0.05 | |
| Sinnott, 2006[51] | Weight based calculation (WBC)<br>n = 128 | Osteopenia: 39%<br>Osteoporosis: 7% | T-score ≤ -2.5 at the hip<br>(T-scores were calculated using the manufacturer's reference values [young Caucasian male database]) | AUC/ROC 0.75 (0.57 to 0.92) | At risk (QUADAS) |
| | BMI<br>n = 128 | | T-score ≤ -2.5 at the hip<br>(T-scores were calculated using the manufacturer's reference values [young Caucasian male database]) | AUC/ROC 0.67 (0.47 to 0.87) | |
| Papaleontiou, 2019[70] | Thyroid cancer<br>n = 539 | n = 539, 6.2% male patients with thyroid cancer;<br>n = 349, 4.0% male patients without thyroid cancer and not on LT4 | Osteoporosis by ICD-9 codes (733.0x) | OR 1.46 (1.26 to 1.68) | Unclear |
| | Thyroid cancer<br>n = 212 | n = 212, 2.4% fractures in male patients with thyroid cancer; | Fracture by ICD and CPT codes | OR 0.70 (0.58 to 0.85) | |

| Author, Year | Risk factors<br>Total n | Prevalence OR | Outcome metric | Effect size (95% CI) | ROB |
|---|---|---|---|---|---|
| | | n = 241, 2.8% fractures in male patients without thyroid cancer | | | |
| Khan, 2013[74] | Smoking<br>n = 2708 | 2239/34665 with osteoporosis, 6.5% of total cohort<br>1506/34665 osteopenic, 4.3% of total cohort<br>588/34665 fragility fracture, 1.7% of total cohort. | Low BMD (Combines osteopenia [T-score ≤ -1.5] and osteoporosis [T-score ≤ -2.5]) | OR 1.2 (1.1 to 1.4) | High risk |
| | Hyperparathyroidism (primary/ secondary),<br>n = 414 | | Low BMD (Combines osteopenia [T-score ≤ -1.5] and osteoporosis [T-score ≤ -2.5]) | OR 2.8 (2.2 to 3.5) | |
| | Hypogonadism<br>n = 1386 | | Low BMD (Combines osteopenia [T-score ≤ -1.5] and osteoporosis [T-score ≤ -2.5]) | OR 2.3 (2.0 to 2.7) | |
| | Obesity<br>n = 10,718 | | Low BMD (Combines osteopenia [T-score ≤ -1.5] and osteoporosis [T-score ≤ -2.5]) | OR 0.8 (0.7 to 0.9) | |
| | Alcoholism<br>n = 3644 | | Low BMD (Combines osteopenia [T-score ≤ -1.5] and osteoporosis [T-score ≤ -2.5]) | OR 0.9 (0.8 to 1.0) | |
| | Malnutrition<br>n = 316 | | Low BMD (Combines osteopenia [T-score ≤ -1.5] and osteoporosis [T-score ≤ -2.5]) | OR 2.0 (1.5, to 2.7) | |
| | Vitamin D deficiency<br>n = 2053 | | Low BMD (Combines osteopenia [T- | OR 2.9 (2.6 to 3.3) | |

| Author, Year | Risk factors<br>Total n | Prevalence OR | Outcome metric | Effect size (95% CI) | ROB |
|---|---|---|---|---|---|
| | | | score ≤ -1.5] and osteoporosis [T-score ≤ -2.5]) | | |
| | Prednisone cumulative dose decile >11136 (mg)<br>n = 1076 | | Low BMD (Combines osteopenia [T-score ≤ -1.5] and osteoporosis [T-score ≤ -2.5]) | OR 8.9 (7.8 to 10.2) | |
| | Smoking<br>n = 2708 (patients with ulcerative colitis) | | Fragility fracture | OR 1.6 (1.2 to 2.0) | |
| | Hyperparathyroidism (primary/ secondary),<br>n = 414 | | Fragility fracture | OR 1.0 (0.6 to 1.7) | |
| | Hypogonadism<br>n = 1386 | | Fragility fracture | OR 1.3 (0.9 to 1.8) | |
| | Obesity<br>n = 10718 | | Fragility fracture | OR 1.2 (1.0 to 1.5) | |
| | Alcoholism<br>n = 3644 | | Fragility fracture | OR 1.8 (1.4 to 2.3) | |
| | Malnutrition<br>n = 316 | | Fragility fracture | OR 2.0 (1.2 to 3.4) | |
| | Vitamin D deficiency<br>n = 2053 | | Fragility fracture | OR 1.9 (1.5 to 2.4) | |
| | Prednisone cumulative dose decile >11136 (mg)<br>n = 1076 | | Fragility fracture | OR 1.8 (1.3 to 2.5) | |
| Weaver 2019[72] | Antipsychotic use<br>n = 5824 | 578 men with antipsychotic use with fractures/5824 men with | Fracture | NR | High risk |

| Author, Year | Risk factors<br><br>Total n | Prevalence OR | Outcome metric | Effect size (95% CI) | ROB |
|---|---|---|---|---|---|
| | | antipsychotic use, 578 men without antipsychotic use with fractures/5667 men without antipsychotic use (approx. 10% in both groups) | | | |

[a] Fine and Gray odds ratio adjusted for death as competing event.

## QUALITY OF EVIDENCE FOR KEY QUESTION 2

### *Studies Examining Osteoporosis and/or Fracture Risk Prediction Tools*

Using the QUADAS tool, there were 2 low risk of bias[28,48] and 6 "at risk of bias" studies across all studies that examined risk assessment tools for osteoporosis and/or fracture among Veterans (Figures 12, 13).[26,51,53,54,56,67] Patterns that led to judgments of "at risk of bias" included: 1) selection of patients (5 studies), 2) interpretation of the index test (5 studies), 3) interpretation of reference standard (1 study), and 4) patient flow (1 study).

### *Studies Examining Individual Risk Factors for Osteoporosis and Fracture*

All studies examining individual risk factors for low BMD or fracture were unclear or high risk of bias (Figures 14, 15, 16).[51,69-79] As a result, the quality of evidence for individual risk factors among male Veterans is very low, and additional work is needed fill these knowledge gaps.

**Figure 12. Risk of Bias for Included KQ 2 Studies Evaluating Combined Risk Factors for Osteoporosis and/or Fracture in Male Veterans**

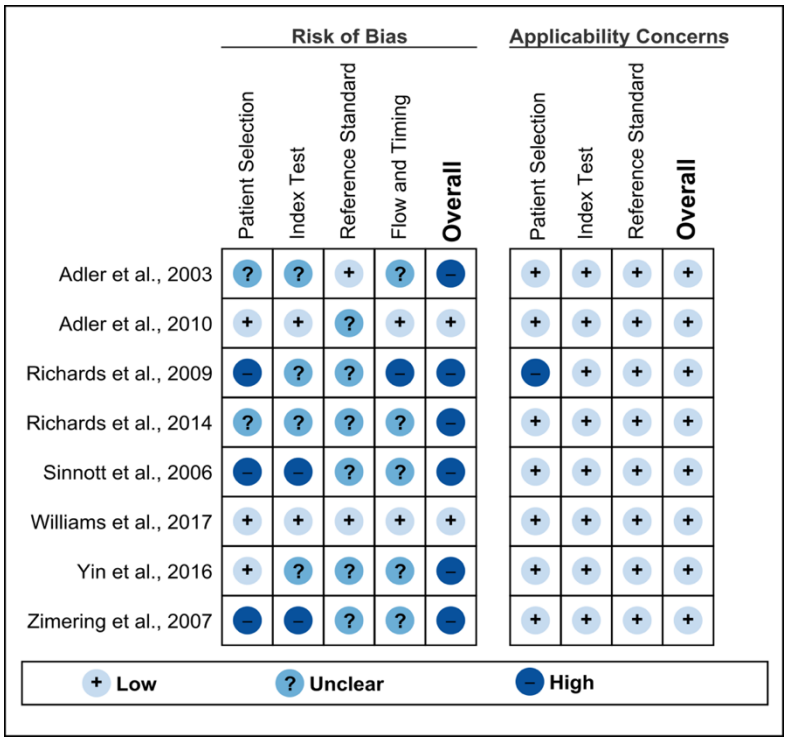
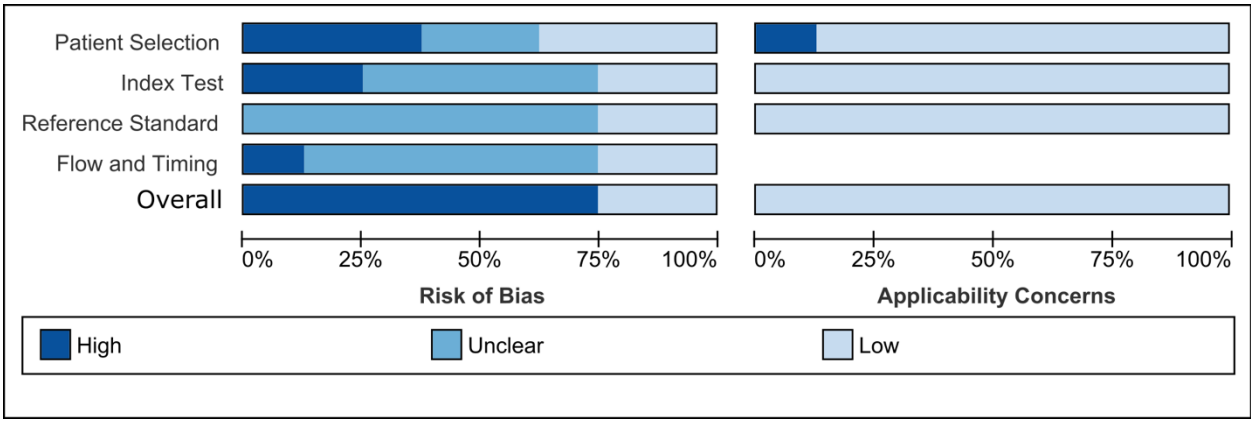
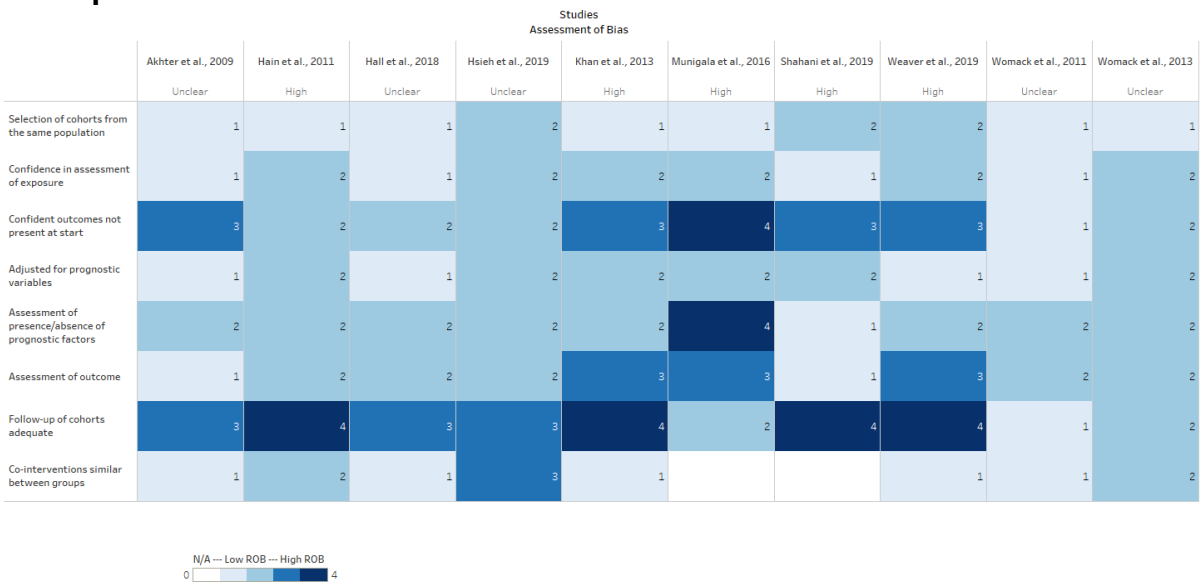



**Figure 13. Risk of Bias Across KQ 2 Studies Evaluating Combined Risk Factors for Osteoporosis and/or Fracture in Male Veterans**



## Figure 14. Risk of Bias for Included Studies Evaluating Individual Risk Factors for Osteoporosis and/or Fracture in Male Veterans




## Figure 15. Risk of Bias for Included Case-Control Study Evaluating Individual Risk Factors for Osteoporosis and/or Fracture in Male Veterans

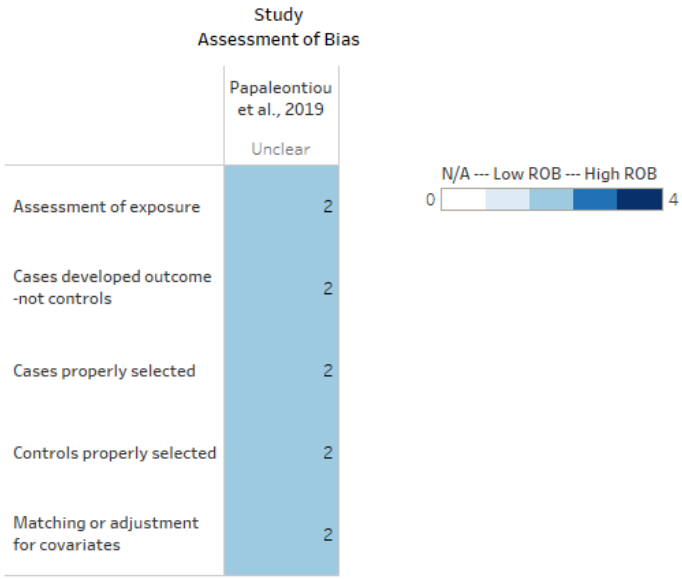
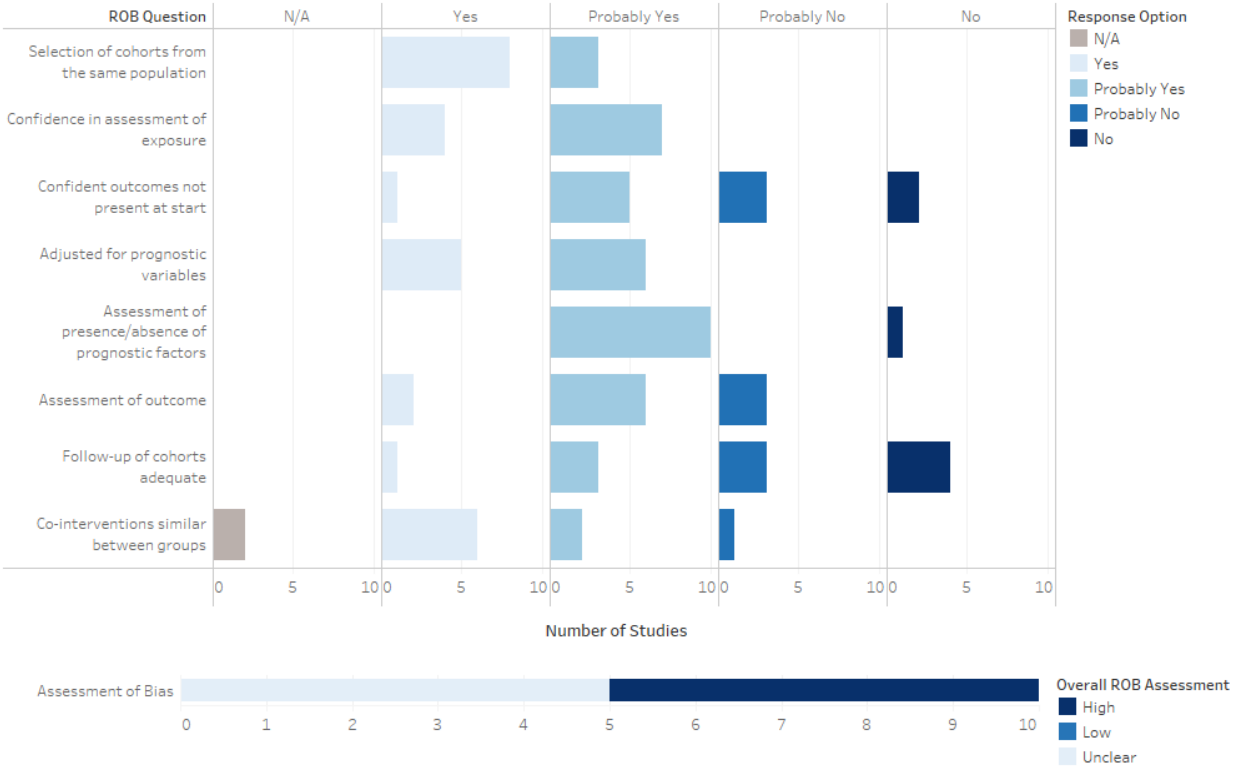


**Figure 16. Risk of Bias Across Included KQ 2 Studies Evaluating Individual Risk Factors for Osteoporosis and/or Fracture in Male Veterans**



# KEY QUESTION 3: What system-level interventions improve uptake of osteoporosis screening among people not identified by a history of low-trauma fracture?

## Characteristics of Included Studies

In total, 20 studies were included examining system-level interventions to improve the uptake of osteoporosis screening among people without a history of low-trauma fracture. Because some studies had more than 1 active intervention arm, a total of 24 intervention arms are described across the 20 studies. Interventions for these studies fell into 8 different categories: 1) provider education (5 studies[92-96]), 2) provider and patient education (3 studies[93,97,98]), 3) provider-focused reminders (4 studies[99-102]), 4) clinical decision support tools (1 study[103]), 5) patient navigation (2 studies[97,104]), 6) patient risk assessment (3 studies[105-107]), 7) patient self-referral (4 studies[108-111]), and 8) patient-focused reminders (1 study[102]). Table 2 shows the evidence profile for the studies.

For KQ 3, we present the detailed results ordered by provider- and patient-focused intervention approaches. When studies assessed the impact of multiple approaches, we synthesized results based on the major emphasis of the intervention approach (*eg*, robust provider education + generic patient education categorized as "provider education"). Details on study characteristics are in Appendix C and a summary of study results by intervention approach are in Tables 6-13.

## Key Points

- Overall, a majority of the identified systems-level interventions in the literature excluded male patients and the most common category of interventions targeted providers (12 studies).
- Provider-focused approaches have mixed effectiveness in improving uptake of osteoporosis screening. Combining provider interventions with targeted patient education improves impact of the intervention but gains are modest significant.
    - Provider education-only interventions (*eg*, CME) demonstrated no increases uptake of osteoporosis screening (4 studies).
    - Provider-focused reminder systems (4 studies) improve uptake of osteoporosis screening via DXA. The impact is greater if provider prompts are coupled with patient education approaches.
    - Clinical decision support tools that combine tailored risk-based education for patients and tailored provider recommendations at the point of clinic visit show promise but have only been evaluated in 1 study.
- Ten studies evaluated the effect of patient-focused approaches on uptake of osteoporosis screening. Overall, patient-focused approaches of patient navigation (2 studies), patient risk assessment (2 studies), patient reminders (1 study), and self-referral systems (4 studies) improve osteoporosis screening via DXA.
    - Coupling patient approaches with provider-focused approaches only marginally increased effectiveness when compared to usual care. Systems redesign approaches that allow patients to self-refer for screening may be more effective if

using fixed appointments compared to open invitations to self-refer without a fixed appointment.

## Detailed Findings

### Provider-focused Intervention Approaches

Twelve studies (8 randomized trials,[92-95,97,98,100,102] 3 nonrandomized trials,[96,99,101] 1 interrupted time series[103]) evaluated the effect of provider-focused approaches on uptake of osteoporosis screening. All studies compared uptake of BMD (*ie*, DXA screening rates) between different provider-focused approaches compared to usual care. Next, we synthesize findings by typology of intervention approach. When meta-analysis was not able to be performed, we computed odds ratios from data reported and display these in the forest plots to add comparability between studies when feasible. We also include estimates provided in the included study.

#### *Provider Education*

Five studies (4 randomized trials[92-95] and 1 nonrandomized trial[96]) focused on provider education as the only intervention approach to improve uptake of BMD testing for osteoporosis screening. Approaches ranged in intensity and dose from a 1-hour case-based session[92] and testing to provider education coupled with a population health management approach[95] (*ie*, list of patients due for BMD testing). Two studies targeted patients age-eligible for osteoporosis screening but 1 only focused on women.[92] Two studies focused on special populations of patients: 1 focused on patients with a history of long-term glucocorticoid use[94] and another on patients with rheumatoid arthritis.[95] Four studies provided sufficient information for meta-analysis and all compared some form of provider education to usual care.[92-95]

Overall, these 4 randomized studies showed no benefit of provider education on uptake of osteoporosis screening (OR 0.98; 95% CI 0.39 to 2.50), although there was significant heterogeneity in intervention effects across studies (Q 8.8; p=0.032; $I^2$ 66.0%) (Figure 15). We conducted exploratory subgroup analysis to assess the impact of enhanced provider education (*ie*, education + audit and feedback to providers, education + list of eligible patients) and special populations used in 2 studies.[94,95] Subgroup analysis by type of provider education and by patient population did little to explain the heterogeneity across studies.

All studies in the meta-analysis were judged at unclear or high ROB. One additional high ROB study also assessed the impact of provider education but was not able to be included in the analysis.[96] This study assessed the impact of didactic lectures on osteoporosis management, compared to no education among primary care physicians, and found no evidence of impact on ordering BMD for eligible patients (33.6% vs 34%).

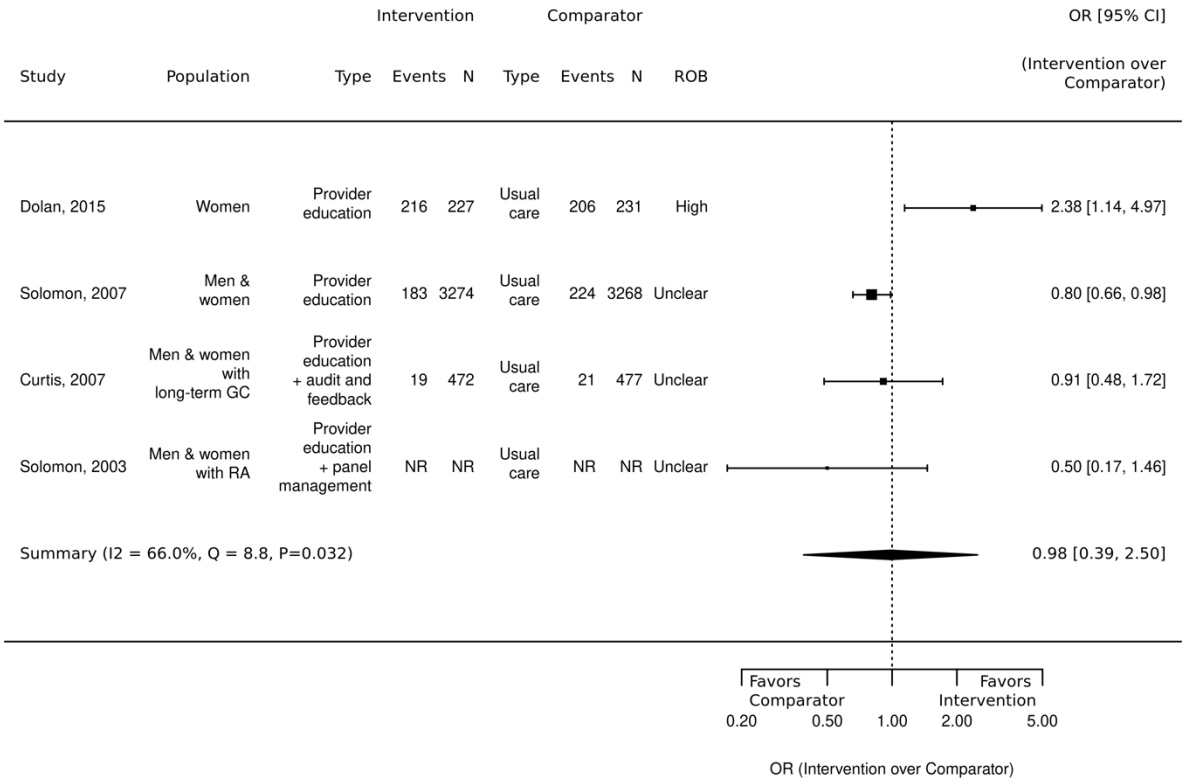
**Table 6. Provider Education**

| Study | Intervention | Results |
|---|---|---|
| Dolan, 2015[92] | Provider education of 1-hour brief case-based session and assessment by 25 multiple choice questions | Provider Education: 216 / 227<br>Usual Care: 206 / 231<br>% difference 0.06% (95% CI 0.01 to 0.11) |

| Study | Intervention | Results |
|---|---|---|
| | | Calculated OR[a] 2.38 (95% CI 1.14 to 4.97) |
| Solomon, 2007[93] | Provider education with academic detailing | Provider Education: 183 / 3,274<br>Usual Care: 224 / 3,268<br>Calculated OR[a] 0.80 (95% CI 0.66 to 0.98) |
| Curtis, 2007[94] | Provider education with audit and feedback<br>Special population: long-term glucocorticoid use | Provider Education: 19 / 472<br>Usual Care: 21 / 477<br>Risk difference: -2 (95% CI -8 to 4)<br>Calculated OR[a] 0.91 (95% CI 0.48 to 1.72) |
| Solomon, 2003[95] | Provider education (mailing of recommendations), academic detailing (dinner meeting), and population health management (curated list of RA patients who were due for BMD testing) vs UC (population: RA) | Provider education and panel management: NR<br>Usual Care: NR<br>OR 0.50 (95% CI 0.2 to 1.5) |
| Pazirandeh, 2002[96] | Provider education (CME) and patient education was conducted for all patients in the study | Provider education (CME): 33.6%<br>Usual Care: 34%<br>OR not reported or calculated |

[a]OR calculated by investigators to create a standard metric across studies.

## Figure 17. Impact of Provider Education on Update of Osteoporosis Screening



## Provider and Patient Education

Three studies (2 randomized controlled trials and 1 cluster-randomized control trial [CRT]),[93,97,98] all with unclear ROB, evaluated the combined effect of provider and patient education on screening rates for osteoporosis to standard care. One of the studies included only men on androgen deprivation therapy,[97] and the other 2 studies included mostly women (92%–97%). Due to differences in study design, interventions assessed, and outcome reporting, a meta-analysis was not performed. Figure 7 displays the estimates for each study.

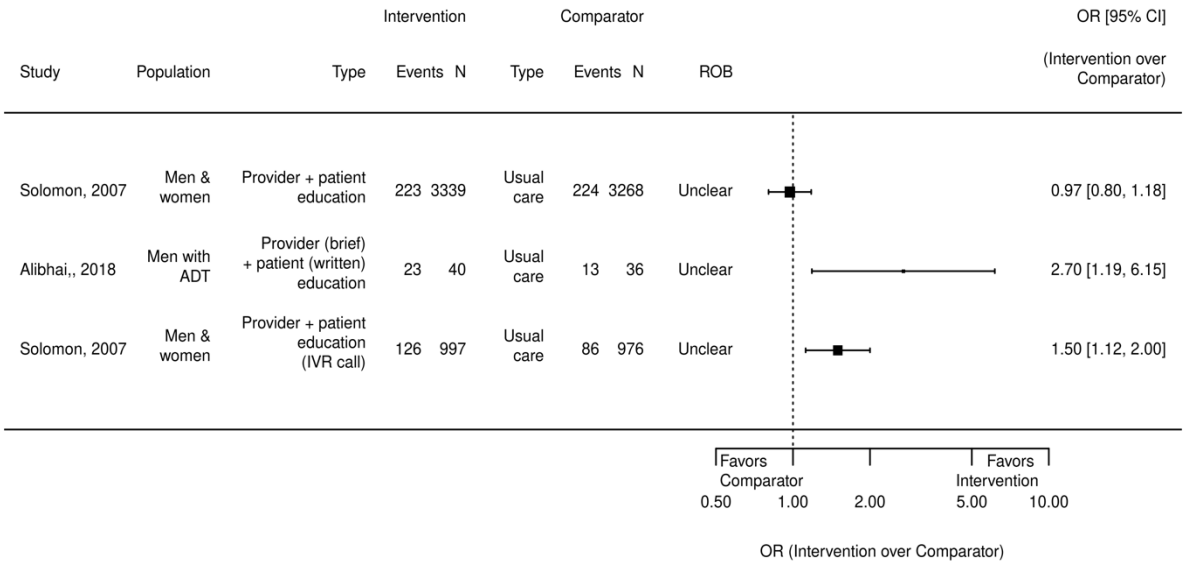
The first study assessed the impact of 1-on-1 education by a physician educator coupled with written CME education and curated list of patients at risk of osteoporosis (n = 1973).[98] Patients received education consisting of an automated telephone call employing interactive voice response (IVR) to provide targeted education and invitation to screen, with the option to transfer the call to schedule BMD testing. At 10 months' follow-up, patients in the intervention arm were more likely to attend BMD screening compared to patients receiving usual care (RR=1.48; 95% CI 1.08 to 2.04). The next study included only males on androgen deprivation therapy (n = 119) and compared usual care with a pamphlet education for patients and printed material to the family physician.[97] At 10 months' follow-up, men exposed to the intervention were significantly more likely to seek a BMD test (OR 2.7; 95% CI 1.19 to 6.15). The last study included 13,455 patients and compared usual care to 3 interventions: generic written patient education, brief provider education through academic detailing by a trained physician educator, and a combined approach of these 2 interventions.[93] Compared to usual care, the combined intervention found no significant impact on uptake of BMD testing at 16 months follow-up (6.9 vs 10.9, p value not reported). Overall, the 2 studies that coupled provider education with targeted patient education demonstrated a statistically significant, modest impact of the combined intervention.[97,98] The study that combined provider education with generic written patient education showed no statistically significant increase in BMD.[93]

### Table 7. Provider and Patient Education

| Study | Intervention | Results |
|---|---|---|
| Solomon, 2007[98] | Patient education of IVR call with education and the option to transfer to scheduling for BMD testing + provider education of individual session with physician educator pharmacist, written education on osteoporosis diagnosis, management and treatment for CME, and curated list of patients in their practice at risk for osteoporosis | Provider Education and patient education: 126 / 997<br>Usual Care: 86 / 976<br>Risk Ratio: 1.48 (95% CI 1.08 to 2.04)<br>Calculated OR[a] 1.50 (95% CI 1.12 to 2.00) |
| Alibhai, 2018[97] | Patient received of 10-page educational pamphlet + mailed education material to family physician<br><br>Special population: men with history of androgen deprivation therapy | Patient Education + Provider Education: 23/40<br>Usual Care: 13/36<br>OR 2.7 (95% CI 1.19 to 6.15)<br>Calculated OR[a] 2.39 (95% CI 0.95 to 6.04) |
| Solomon, 2007[93] | Patient education (not tailored) + in-person provider education by a physician educator informed by academic detailing principles | Provider Education and patient education: 223 / 3,339<br>Usual Care: 224 / 3,268<br>Calculated OR[a] 0.97 (95% CI 0.80 to 1.18) |

[a]OR calculated by investigators to create a standard metric across studies.

**Figure 18. Impact of Provider Education plus Patient Education on Uptake of Osteoporosis Screening**




## Provider-focused Reminders

Four studies (2 nonrandomized trials,[99,101] 2 CRTs[100,102]) evaluated the impact of provider-focused reminders on uptake of BMD testing. Three of these studies were of unclear ROB[100-102] and 1 had high ROB.[99] All but 1 study used EMR-based provider reminders.[100] Two studies assessed the impact of additional intervention approaches of patient education[100] or panel management to provider reminder systems.[101] Only 1 study included men; all other studies focused on women age-eligible for BMD screening.[99] Studies were too varied to conduct meta-analysis. Figure 8 displays point estimates for each of the arms in 3 studies; we were not able to convert 1 study to a common metric for visual display.[102]

The first CRT was a large (n = 10354) assessment of a generic mailed patient reminder alone or paired with physician EHR-based prompts for osteoporosis screening conducted in 15 primary care clinics.[102] Patient reminders alone (24.1%), or in combination with provider prompts (28.9%), significantly increased BMD testing compared to usual care (10.8%; $p < 0.001$). The next CRT conducted an assessment of chart reminders in the form of paper sticky notes alone or in combination with mailed generic patient education in 5 primary care clinics (n = 195).[100] Chart reminders in combination with patient education (45.2%; OR 5.47, $p = 0.029$) increased BMD testing among age-eligible women compared to usual care (9.7%). Yet, while chart reminders alone increased screening, the impact was not statistically significant (OR 2.37, $p = 0.156$).

The first nonrandomized study assessed the impact of EHR-based reminders for 4 preventive practice behaviors (*ie*, osteoporosis screening via BMD, influenza and pneumococcal vaccinations, and health care proxy designation) alone or in combination with a panel manager compared to usual care (n = 4660).[101] The panel manager approach was an off-site administrative assistant who reviewed patient panels for patients due for any of the 4 behaviors. The panel

manager then emailed the provider a list of patients due and asked permission to contact. If permission was granted, the panel manager called the patient and facilitated completion of the needed services. Compared to usual care, both arms increased BMD testing, but only the provider reminder arm combined with panel management was statistically significant (OR 2.31; 95% CI 1.55 to 3.43). The next nonrandomized study (high ROB) also assessed the impact of EMR-based reminder on a cluster of 3 practice behaviors but included actionable links to computerized orders (n = 3849).[99] While provider reminders increased screening, the impact was not significant (OR 1.29; 95% CI 0.82 to 2.02). Overall, the 4 included studies suggest that provider reminders improve uptake of osteoporosis screening via BMD. The impact is greater if provider prompts are coupled with patient education approaches.

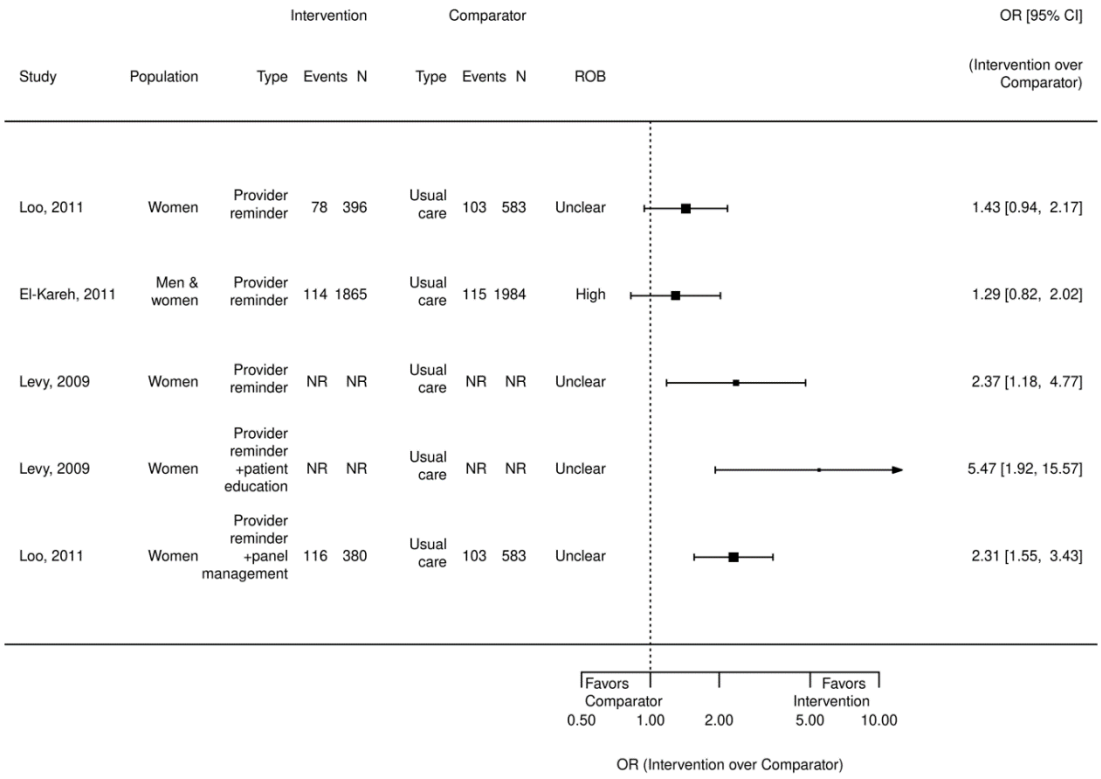
**Table 8. Provider-focused Reminders**

| Study | Intervention | Results |
|---|---|---|
| El-Kareh, 2011[99] | EMR-based reminders with actionable link to computerized order for 3 practice behaviors, including osteoporosis screening | System reminder - provider: 114 / 1,865<br>Usual Care: 115 / 1,984<br>OR 1.29 (95% CI 0.82 to 2.02) |
| Levy, 2009[100] | Provider: Printed sticky practices could place on patient charts<br><br>Patient: Mailed generic patient education | Chart reminder + mailed patient education vs Usual Care:<br>OR 5.47 (p = .029)<br>Calculated[a] (95% CI 1.92 to 15.57)<br><br>Chart reminder vs Usual Care: OR 2.37 (p = .156)<br>Calculated[a] (95% CI 1.18 to 4.77) |
| Loo, 2011[101] | EMR reminders for 4 preventive practice behaviors, including osteoporosis screening<br>EMR + panel manager arm off-site panel manager reviewed, geriatrics patient list for patients due for any of 4 behaviors, emailed provider for permission to contact patient and called patient to facilitate completion of needed health services.<br><br>If unable to contact patient, a letter with the same content that would have been provided by telephone was sent | Provider reminder vs Usual Care:<br>OR 1.43 (95% CI 0.94 to 2.17)<br><br>Provider reminder + panel management vs usual care<br>OR 2.31 (95% CI 1.55 to 3.43) |
| Lafata, 2007[102] | EHR-based provider reminder + generic mailed patient reminder | All patients 65+:<br>Usual Care (10.8%) vs Patient mailed reminder and physician prompt (28.9%), (p<0.001).<br><br>Age 65: Usual Care 17.0 (13.8 to 20.9), Patient mailed reminder and physician prompt: 30.3 (27.8 to 32.9)<br>Age 75: Usual Care 10.1 (8.0 to 12.6) Patient mailed reminder and physician prompt: 27.0 (24.7 to 29.4)<br>Age 85: Usual Care: 5.8 (4.5 to 7.3) |

| Study | Intervention | Results |
|---|---|---|
| | | Patient mailed reminder and physician prompt: 23.9 (21.8 to 26.2)<br><br>BMD Testing Age x Patient mailed reminder and physician prompt: Beta = 0.04 (SE = 0.01), p < 0.01<br><br>OR not reported and unable to be calculated |

[a]OR calculated by investigators to create a standard metric across studies.

## Figure 19. Impact of Provider-focused Reminders on Uptake of Osteoporosis Screening




## Clinical Decision Support Tools

One high ROB study assessed the impact of a clinical decision support tool on age-eligible men (aged 65+) and women (aged 50+) using an interrupted time-series design.[103] The study was conducted in 3 family practices encompassing 5 physicians and 2840 eligible patients. The intervention was completed at the time of the clinical visit with 3 components: risk assessment questionnaire for patients completed on a tablet, paper-based best practices prompt for physicians based on responses from the patient, and customized osteoporosis educational sheet given to patients at the end of the physician visit summarizing individualized risk and

suggestions for managing those risks. Results showed an increase from baseline in initiation of screening (3.4%; $p < 0.001$).

**Table 9. Clinical Decision Support Tool**

| Study | Intervention | Results |
|---|---|---|
| Kastner, 2014[103] | Risk assessment tool given to patient in waiting room. The system gives a individualized, person specific printout for provider with treatment recommendations. The system also gives the patient a specific educational printout at the end of the visit. | Clinical decision support tool<br>Mean 6.15 (SD 2.24)<br><br>Usual Care<br>Mean 2.79 (SD 1.27)<br><br>% change: 3.4% (95% CI 2.03 to 4.68)<br><br>OR not reported or calculated |

## Patient-focused Intervention Approaches

Ten studies (8 randomized trials,[97,102,105-109,111] 1 controlled before and after study,[110] 1 interrupted time-series study[104]) evaluated the effect of patient-focused approaches on uptake of osteoporosis screening. All studies compared uptake of BMD (*ie*, DXA screening rates) between different patient-focused approaches or compared to usual care. Below we synthesize findings by typology of intervention approach. When meta-analysis was not able to be performed, we computed odds ratios from data reported and display these in the forest plots to add comparability between studies when feasible. We also include the estimates provided in the study.

### *Patient Navigation*

Two studies (1 RCT,[97] 1 interrupted time-series study[104]) assessed the impact of patient navigation on uptake of osteoporosis screening. Both studies were judged to be unclear ROB. The first study was a single-center RCT with men on androgen deprivation therapy (n = 119) and compared usual care with a pamphlet education for patients plus individual patient navigation by a bone health care coordinator who followed up at least twice over 3 months to facilitate BMD ordering.[97] At 10 months' follow-up, men exposed to the intervention were significantly more likely to seek a BMD test (OR 2.7; 95% CI 1.19 to 6.15). The interrupted time-series study was conducted in women only (pre-intervention n = 1782; post-intervention n = 1981) and assessed the impact intervention consisting of mailed and telephone contacts with an outreach coordinator authorized to schedule a screening appointment without a provider visit first compared to usual care.[104] Over 13 months of the intervention period, the percent increase in the population screened that was attributable to the outreach coordinator was 13% ($p < 0.001$). Overall, both studies suggest a positive impact on patient navigation strategies on uptake of osteoporosis screening.

**Table 10. Patient Navigation**

| Study | Intervention | Results |
|---|---|---|
| Alibhai, 2018[97] | Patient received 10-page educational pamphlet + 1-on-1 session with a bone health care coordinator who followed up at least twice over 3 months to facilitate BMD ordering | Patient education + bone health care coordinator navigation : 23/40<br>Usual Care: 13/36<br>OR: 2.7 (95% CI 1.19, 6.15) |
| Denberg, 2019 [104] | An invitation letter was mailed to each patient and included the name of the patient's PCP, summarized the USPSTF recommendations and rationale for DXA, and encouraged the patient to contact the call center to arrange an examination. If the patient did not respond within 2 weeks, the outreach coordinator made up to 3 calls to the patient's home at different times of the day over a period of 8 weeks, leaving a voice message on the first attempt. | Patient education + navigation vs usual care: p < .001<br><br>OR: not reported or calculated |

### *Patient Risk Assessment*

Three studies (1 CRT, 2 individual-level RCTs) assessed the impact of patient risk assessment (*eg*, calculating FRAX).[105-107] Two of these studies were judged unclear ROB,[105,107] and 1 low ROB[106]; only 1 study included men.[106] Due to differences in study design, interventions assessed, and outcome reporting, we did not perform meta-analysis.

The first study was a low ROB RCT that assessed the impact of a community pharmacist screening program on osteoporosis testing via BMD and treatment among 262 people age-eligible for screening or with at least 1 major risk factor for fracture.[106] The intervention consisted of generic printed patient education, quantitative heel ultrasound, and risk assessment feedback via the community pharmacist compared to a usual care. At 4 months' follow-up, community dwelling participants exposed to the intervention were significantly more likely to undergo BMD testing compared to controls (RR = 2.20; 95% CI 1.2 to 4.1). The next study was an unclear ROB RCT that sent 34,229 randomized women in Denmark to an intervention or control group.[107] All women in the study were mailed a FRAX and asked to mail it back. Women in the intervention arm were offered a DXA scan if their 10-year probability of MOF was 15% or greater. Women in the control group received no further communication. Approximately 80% returned the questionnaire; 10-year risk of osteoporotic fracture was able to be calculated for 61% of all women in the study. The intervention had no overall impact of MOF after a mean follow-up time of 5 years (p=0.682). The last study was an unclear ROB RCT among 4685 women aged 50 to 64 with an elevated risk of developing osteoporosis.[105] Women were randomized to usual care or 1 brief telephonic IVR call which calculated a fracture risk-score based on women's responses, history of BMD testing, and intentions to discuss osteoporosis with their physician. The call ended with a recommendation to discuss BMD testing with their physician. Women randomized to receive IVR intervention were significantly more likely to obtain BMD within 12 months compared to usual care (25.6% intervention vs 18.6% in usual care; p < 0.001). Overall, patient risk assessment showed promise at increasing BMD screening

when coupled with individualized feedback on risk in 2 studies.[106] Only 1 study assessed risk of fracture and found no impact of administering the FRAX coupled with invitation to screen and no feedback on risk.[107]

**Table 11. Patient Risk Assessment**

| Study | Intervention | Results |
|---|---|---|
| Heyworth, 2014[105] | Patient IVR was a call lasting 4-5 minutes in which the automated response recognized the patient and confirmed identification. The call ascertained the patients' fracture risk score and the following information: history of BMD testing, plans to follow up with their physician to discuss osteoporosis. | Patient Risk Assessment and feedback: 385 / 1,565<br>Usual Care: 290 / 1,558<br>Percent difference: 6% (p < 0.001) |
| Yuksel, 2010[106] | Risk assessment with community pharmacist via quantitative heel ultrasound, plus generic patient education | Patient Risk Assessment and feedback: 28 / 129<br>Usual Care: 13 / 133<br>Risk Ratio: 2.2 (95% CI 1.2, 4.1) |
| Rubin, 2018[107] | All participants were mailed a questionnaire. Questionnaires with fewer than 3 missing items were used to calculate a FRAX score. Intervention arm participants received an offer for a DXA if they had a 10-year probability of fracture of 15% or more. Results of DXA were shared with patient and GP | Patient Risk Assessment and feedback: 1,697 / 17,072<br>Usual Care: 1,719/17,157<br>Sub-hazard ratio of MOF: 0.986 (95% CI 0.922; 1.055)<br><br>Patient Risk Assessment and feedback: 534 / 17,072<br>Usual Care: 532 /17,157<br>Sub-hazard ratio hip fracture: 1.002 (95% CI 0.889; 1.130) |

### *Patient Self-referral*

Four studies (2 RCTs, 1 CRT, 1 controlled before-after study) evaluated the effect of patients' ability to self-refer for osteoporosis screening.[108-111] All studies were judged unclear ROB. No studies included men. In total, 19,740 women were included in those trials, and screening rates from 90 days to 5 months were evaluated. All interventions across the 4 studies involved a mailed reminder. Three of the studies' interventions were mailed reminders to call and schedule an osteoporosis screening compared to a no-contact control.[108-110] One study compared 3 different types of self-referral invitations (*ie*, fixed appointment that required telephone confirmation of intent to attend screening, fixed appointment with option to change time but no need to call to confirm, open invitation with no preassigned appointment slot)[111] and 1 study compared mailed self-referral alone or with additional patient education in the form of a DVD by trial site (*ie*, Northwestern, Georgia).[109] Due to differences in study design and outcome reporting, meta-analysis could not be performed. Figure 18 displays the point estimates for each of the 9 comparisons across the 4 studies.

In the 3 interventions with mailed invitations to self-referral only compared to usual care, self-referral resulted in a statistically significant increase in osteoporosis screening (OR range: 2.70 to 4.87).[108-110] In the study assessing additional patient education via a mailed DVD, self-referral significantly improved osteoporosis screening over no-contact control at each of the 2 trial sites, but the increases were similar to the condition receiving self-referral invitation alone.[109] In the study that assessed 3 different types of self-referral procedures,[111] fixed appointment invitations (75% screened) and confirmable invitations (69% screened) significantly increased osteoporosis screening over open invitations to call and schedule a screening (54%; $p < 0.0001$). Yet fixed reminders did not result in statistically significant increases in osteoporosis screening over confirmable reminders ($p = 0.083$). Overall, all studies demonstrated a statistically significant impact of self-referral approaches to improving osteoporosis screening.

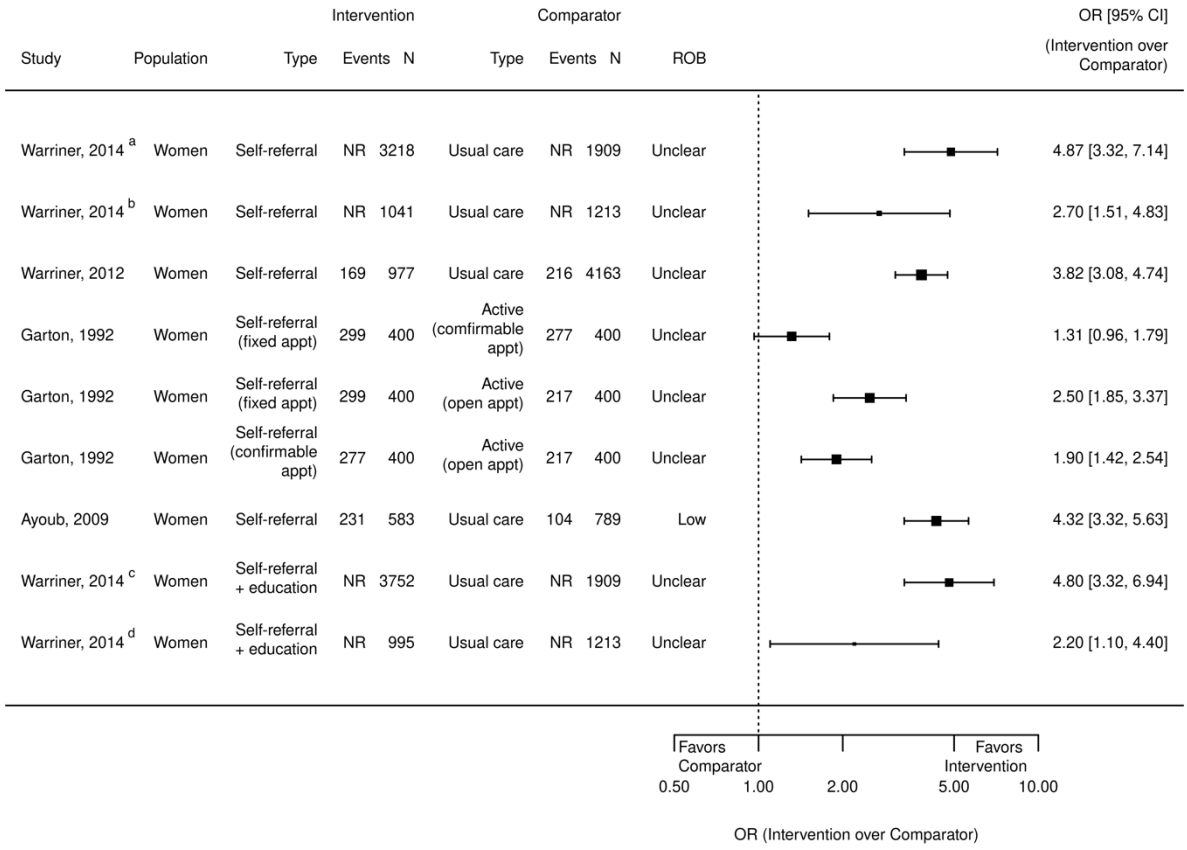
**Table 12. Patient Self-referral**

| Study | Intervention | Results |
|---|---|---|
| Warriner, 2014[109] | Intervention 1: Selected eligible patients received a mailed invitation for self-referral for DXA scan. All physicians were educated and provided information regarding osteoporosis.<br><br>Intervention 2: Patients received a mailed invitation for self-referral for DXA scan and patient education via DVD. All physicians were educated and provided information regarding osteoporosis. | Kaiser Permanente Northwest self-referral vs Usual Care: OR 4.9 (95% CI 3.3, 7.1)<br><br>Kaiser Permanente Northwest self-referral with patient educational materials vs Usual Care: OR 4.8 (95% CI 3.3, 6.9)<br><br>Kaiser Permanente Georgia self-referral vs Usual Care: OR 2.7 (95% CI 1.5, 4.8)<br><br>Kaiser Permanente Georgia self-seferral with educational materials vs Usual Care: OR 2.2 (95% CI 1.1, 4.4) |
| Warriner, 2012[108] | Patients in this group were sent a letter and brochure detailing the importance of osteoporosis screening and offering guidance on how to self-schedule a DXA scan | Patient self-referral: 17.3%<br>Usual Care: 5.2%<br>OR: 2.9 (95% CI 1.7, 4.8)<br>Calculated OR[a]: 3.82 (3.08, 4.74) |
| Ayoub, 2009[110] | Patient self-referral with 1 follow-up call | Patient self-referral: 231/583<br>Usual Care: 104/789<br>Calculated OR[a]: 4.32 (3.32, 5.63) |
| Garton, 1992[111] | Women were randomized to 3 types of mailed reminders:<br><br>Intervention 1: confirmable BMD appointment<br><br>Intervention 2: fixed BMD appointment<br><br>Intervention 3: open invitation to call and schedule a BMD screening but no set appointment | Patient self-referral (Fixed appt) vs open appointment:<br>Risk difference 12 (95% CI 14, 27)<br>Calculated OR[a]: 2.50 (95% CI 1.85, 3.37)<br><br>Patient self-referral (Confirmable appt) vs open appointment:<br>Risk difference 15 (95% CI 8, 22)<br>Calculated OR[a]: 1.90 (95% CI 1.42, 2.54)<br><br>Patient self-referral (Fixed appt) vs Patient self-referral (Confirmable appt):<br>Risk difference 6 (95% CI -1, 12) |

| Study | Intervention | Results |
|---|---|---|
| | | Calculated OR[a]: 1.31 (95% CI 0.96, 1.79) |

[a]OR calculated by investigators to create a standard metric across studies.

## Figure 20. Patient Self-referral on Uptake of Osteoporosis Screening




[a] Kaiser Permanente Northwest Self-referral vs Usual Care
[b] Kaiser Permanente Georgia Self-referral vs Usual Care
[c] Kaiser Permanente Northwest Self-referral with Patient Educational Materials vs Usual Care
[d] Kaiser Permanente Georgia Self-referral with Educational Materials vs Usual Care

### *Patient-focused Reminders*

One unclear ROB CRT conducted in 15 primary care clinics (n = 10,354) assessed an intervention consisting of 2 mailed patient reminders timed to occur with the primary care visit and 1 month after the primary care visit alone, or paired with physician EHR-based prompts for osteoporosis screening.[102] This study focused on women only. Patient reminders alone (24.1%), or in combination with provider prompts (28.9%), significantly increased BMD testing compared to usual care (10.8%; $p < 0.001$).

**Table 13. Patient Reminders**

| Study | Intervention | Results |
|---|---|---|
| Lafata, 2007[102] | Two generic mailed patient reminders sent to women with a recent visit to their primary care physician. | All patients 65+:<br>Patient mailed reminder: 24.1% (unadjusted)<br>Usual Care: 10.8% (unadjusted)<br><br>Results by age brackets:<br>Age 65:<br>Usual Care: 17.0% (13.8, 20.9)<br>Patient Mailed Reminder: 23.2% (20.6, 25.9)<br><br>Age 75:<br>Usual Care: 10.1% (8.0, 12.6)<br>Patient Mailed Reminder: 18.7% (16.5,21.0)<br><br>Age 85:<br>Usual Care: 5.8% (4.5, 7.3)<br>Patient Mailed Reminder: 14.8% (13.1, 16.8)<br><br>BMD Testing Age × Mailed Reminder: Beta = .03 (SE = .01), p = .01.<br><br>OR: not reported or calculated |

## QUALITY OF EVIDENCE FOR KEY QUESTION 3

The ROB for patient-reported outcomes was judged low for 4 studies,[100,104,106,110] unclear for 1 study,[97] and high for 1 study.[96] For objectively reported outcomes, the ROB was judged low for 3 study[104,106,110] and unclear for 13 studies.[93-95,98,100-103,105,107-110]

Patterns that led to judgments of higher ROB included: 1) selection bias related to random sequence generation (n = 9), 2) differences in baseline measurement for study conditions (n = 11), and 3) outcome assessments that did not clearly blind to intervention assignment (n = 9). In addition to the lack of randomization for the 5 nonrandomized studies, (3 nonrandomized trials, 1 controlled before and after, 2 interrupted time-series studies), baseline provider contamination and incomplete outcome data led to a judgement of high ROB. The ROB ratings and assessments for each study are shown in Figures 19, 20, 21, and 22.

**Figure 21. Risk of Bias Ratings for the Included Studies**

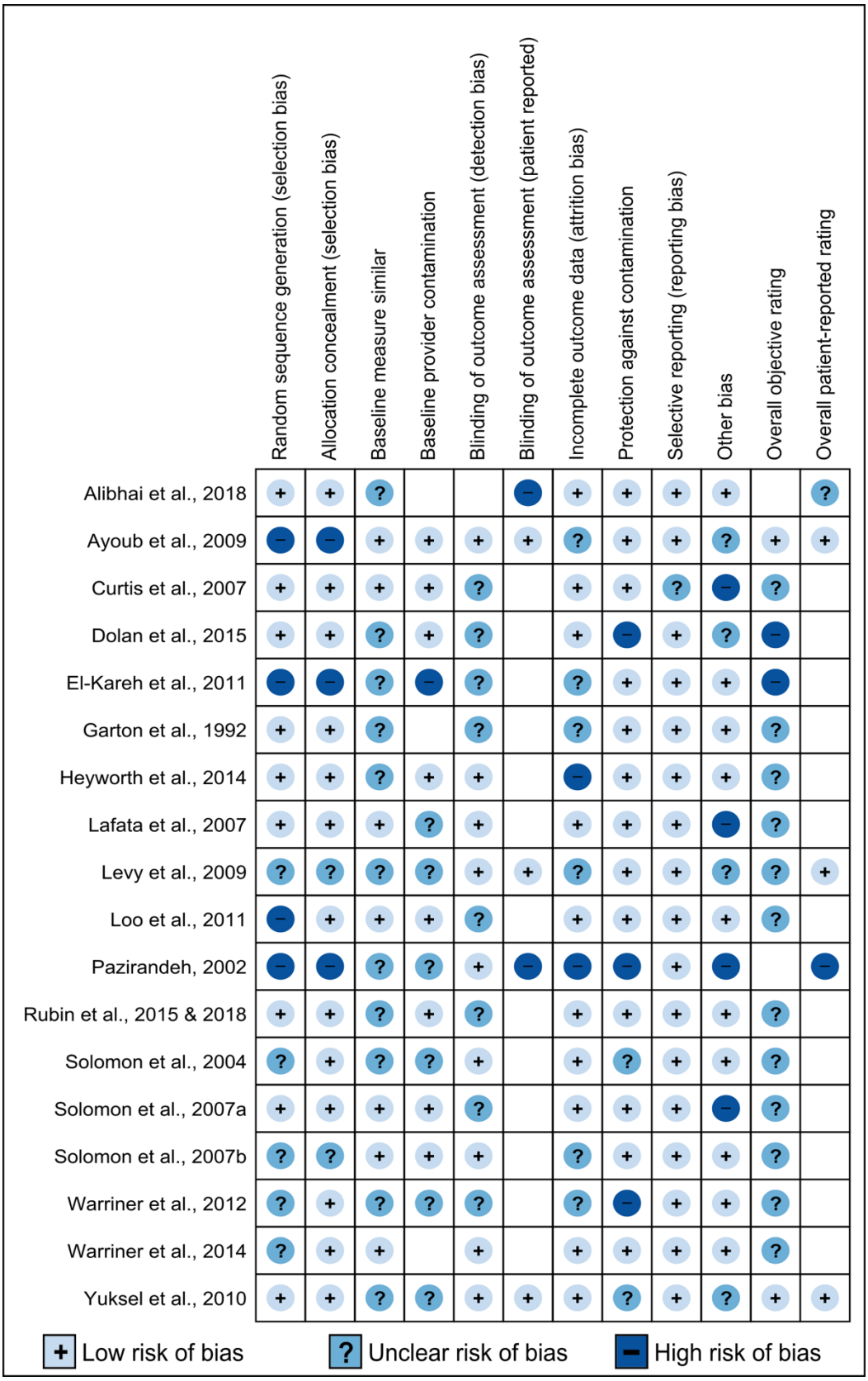


**Figure 22. Risk of Bias Assessment Across Included Studies**

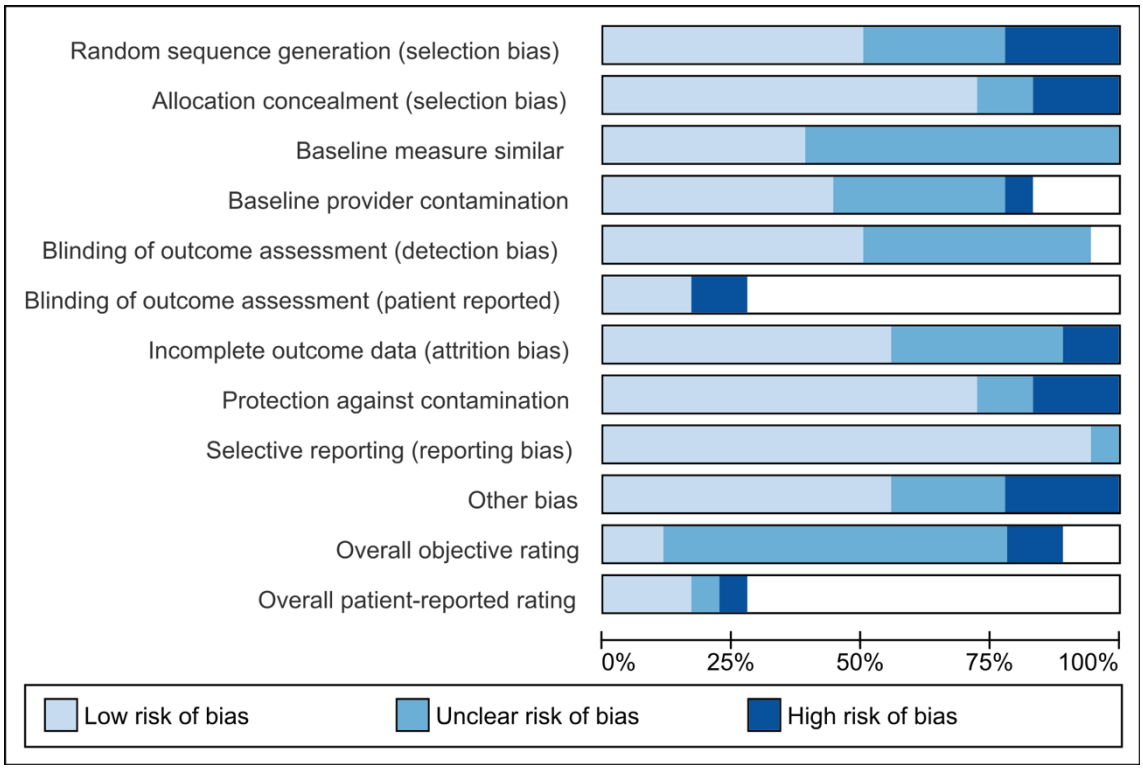



**Figure 23. Risk of Bias Ratings for the Interrupted Time-Series Studies**

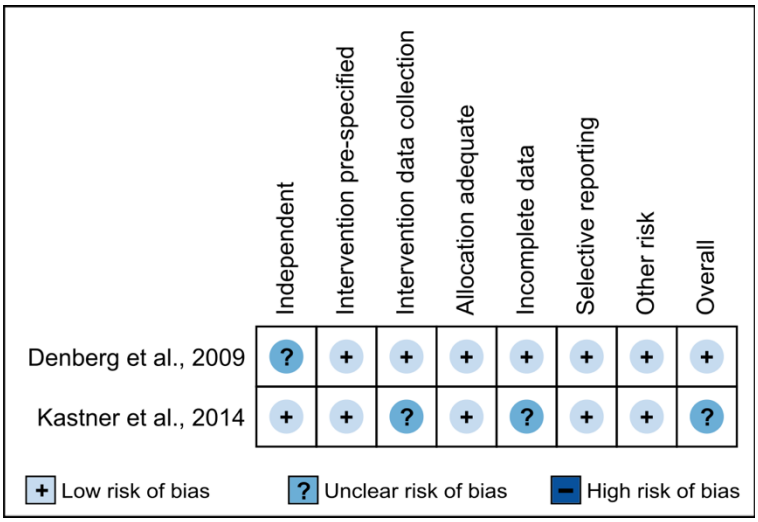
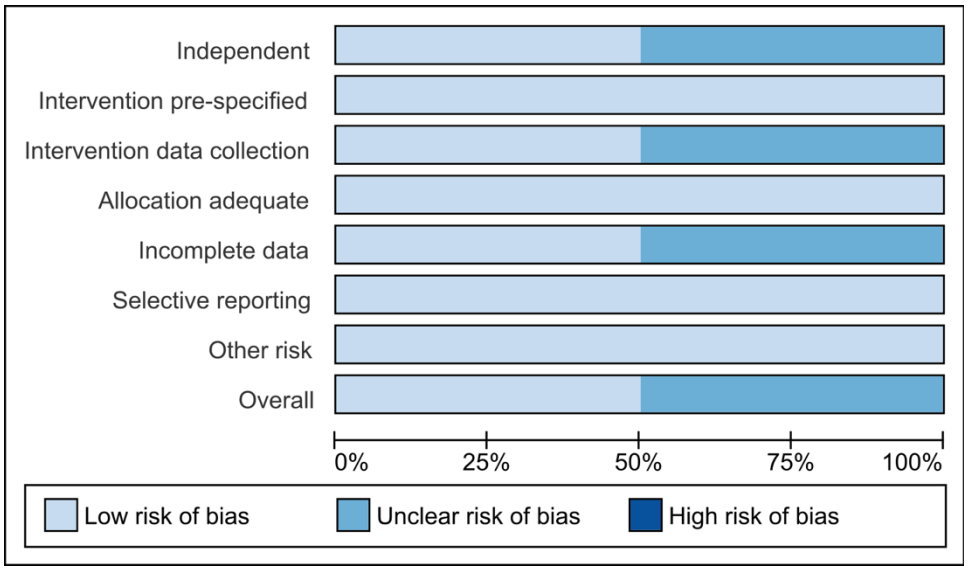


**Figure 24. Risk of Bias Assessment Across the Interrupted Time-Series Studies**



# SUMMARY AND DISCUSSION

Primary prevention of osteoporosis is largely sought through screening to identify those at highest risk of fracture-related morbidity. While screening women aged 65 and older for osteoporosis is standard clinical practice, and is associated with fracture risk reduction,[112] there is uncertainty about the role of screening among men.[3] Further, there is significant uncertainty about *how* to screen men when screening is determined to be warranted. Thus, fracture risk-assessment tools, such as the FRAX® risk assessment tool, have been developed to identify those who are at high risk for osteoporosis and fracture. Screening by fracture risk tool first, rather than by DXA, has been proposed as an alternate means to identify those at increased risk for fracture.[9]

The issue of screening for osteoporosis among men is particularly pertinent to the VHA. Veterans are at higher risk for osteoporotic fractures than non-Veterans.[10]

However, continued research is needed to support optimal choice and use of clinical risk assessment tools specific to male Veterans, including head-to-head comparisons of specific tools to help guide clinical decision-making. From a clinical perspective, risk assessment tools like the OST, Garvan, and MORES may be more likely to be implemented in the clinical setting if the variables are easily accessible via the EHR and still report moderate to excellent discrimination. Given that the VHA is an integrated health system, it is particularly well suited to conducting research on system-level interventions to improve screening for osteoporosis and fracture risk.

## SUMMARY OF EVIDENCE BY KEY QUESTION

### Key Question 1 Summary

**KQ 1:** Among males not identified by a history of low-trauma fracture, is there a clinical risk tool (*eg*, FRAX®) that identifies patients at highest risk of osteoporosis or major osteoporotic fracture?

We sought to synthesize the evidence on clinical risk assessment tools that identify men at highest risk of osteoporosis or major osteoporotic fracture. Overall, we identified 37 studies that met our inclusion criteria encompassing 18 different clinical risk tools. There were 19 studies that assessed the FRAX risk assessment tool or a modified version of the FRAX risk assessment tool. Tools varied in their complexity, ranging from only 2 risk factors (*ie*, OST) to more than 20 risk factors (*ie*, QFracture). Nine studies assessed the OST/OSTA. QFracture (2 different versions) was used in 5 studies. Four studies evaluated the MORES risk tool. The Garvan tool was assessed in 4 studies. The remaining tools were each evaluated in only 1 study. The FRAX tool was the only tool assessed with all 3 outcomes: MOF, hip fracture, and osteoporosis. While we aimed to synthesize all relevant risk assessment metrics, sensitivity and specificity were underreported across studies, and included studies often did not report the necessary elements to compute these outcomes. Thus, we relied on the AUC, as it was the most commonly reported outcome across included studies.

Overall, we found little evidence that directly compared risk assessment tools within the same population of men. Among the risk assessment tools assessed, the OST has good discrimination in predicting osteoporosis by DXA and had two, easily obtainable variables (AUC ranging from 0.632 to 0.890). The other 2 tools identified that predicted risk of osteoporosis (FRAX, MORES)

had high heterogeneity in AUC, ranging from poor to excellent discrimination (AUC ranging from 0.596 to 0.870). Tools such as FRAX, QFracture, and Garvan display poor to excellent discrimination in predicting hip fracture and MOF (0.609 to 0.930 for hip fracture; 0.618 to 0.810 for MOF). Yet, among men not identified via prior fracture, the FRAX risk assessment tool has better discrimination in predicting hip fracture than MOF and osteoporosis diagnosis. It is worth noting that several studies examining osteoporosis as an outcome in men used a male normative database to define low BMD, whereas FRAX uses the female normative database. This may have altered the discrimination of FRAX to identify osteoporosis in men. Limited evidence was identified for use of FRAX in special populations such as individuals with HIV and those on androgen deprivation therapy (ADT), but it was generally found to perform worse among these groups. Less common tools were reported by 9 studies and had variable discriminatory ability.

Risk of developing osteoporosis, hip fracture, or MOF were deemed the outcomes critical to decision-making. Thus, these are the outcomes for which we conducted certainty of evidence (COE) ratings. These COE judgments reflect the degree of confidence we have in our summary findings. For each outcome of interest, we present the COE organized first by outcome modeled (*ie*, osteoporosis, hip fracture, MOF) for each of the major risk assessment tools, and then by sensitivity, specificity, and AUC. COE ratings are summarized below, with supporting information provided in Table 14.

- We found *low to very low* COE to support that OST, FRAX, and MORES identifies men at high risk of osteoporosis across all 3 risk assessment metrics of sensitivity, specificity, and AUC.
- We found *low to very low* COE to support that FRAX identifies men at high risk of hip fracture across all 3 risk assessment metrics of sensitivity, specificity, and AUC. We found the Garvan too was *low* COE for sensitivity and specificity but *moderate* COE using the AUC measure. QFracture was also judged to be *low* COE for hip fracture identification among men as assessed by AUC alone.
- We found *very low* COE that FRAX identifies men at risk for MOF as assessed by AUC only. For QFracture, we found *moderate* COE that this risk assessment tool identifies men at high risk of MOF.
- Of the included studies, 50% were judged to be at high risk of bias using QUADAS-2.

### Table 14. Certainty of Evidence for Main Outcomes of Osteoporosis Risk Assessment Tools

| Outcome | Number of Studies (N) | Findings | Certainty of Evidence (Rationale) |
|---|---|---|---|
| **OSTEOPOROSIS** | | | |
| ***OST*** | | | |
| Sensitivity | 8 studies (6,805 participants) | Sensitivity range: 0.688 to 0.930 (range of 95% CI 0.588 to 0.997) | **Low certainty that OST identifies men at high risk of osteoporosis** (rated down for very serious risk of bias) |
| | 1 study | 0.64 (95% CI NR) | |

| Outcome | Number of Studies (N) | Findings | Certainty of Evidence (Rationale) |
|---|---|---|---|
| | (282 patients with rheumatoid arthritis) | | |
| Specificity | 8 studies (6,805 participants) | Specificity range: 0.336 to 0.660 (range of 95% CI 0.292 to 0.735) | **Very low certainty that OST identifies men at high risk of osteoporosis** (rated down for very serious risk of bias, indirectness, inconsistency) |
| | 1 study (282 patients with rheumatoid arthritis) | 0.54 (95% CI NR) | |
| AUC | 8 studies (6,805 participants) | AUC range: 0.632 to 0.890 (range of 95% CI 0.535 to 1.030) | **Low certainty that OST identifies men at high risk of osteoporosis** (rated down for very serious risk of bias) |
| ***FRAX*** | | | |
| Sensitivity | 3 studies (6,267 participants) | Sensitivity range: 0.390 to 0.766 (range of 95% CI 0.270 to 0.825) | **Very Low certainty that FRAX identifies men at high risk of osteoporosis** (rated down for very serious risk of bias, very serious inconsistency, and very serious imprecision) |
| | 1 study (400 participants) | 58.3 men 74 or younger<br><br>63.8 men 75 or older | |
| Specificity | 3 studies (6,267 participants) | Specificity range: 0.427 to 0.890 (range of 95% CI 0.384 to 0.910) | **Very low certainty that FRAX identifies men at high risk of osteoporosis** (rated down for very serious risk of bias, serious indirectness, very serious inconsistency, and very serious imprecision) |
| | 1 study (400 participants) | 58.4 (95% CI NR) Men 74 or younger<br><br>65.8 (95% CI NR) Men 75 or older | |
| AUC | 3 studies (6,267 participants) | AUC range 0.596 to 0.790 (range of 95% CI 0.547 to 0.840) | **Very low certainty that FRAX identifies men at high risk of osteoporosis** (rated down for very serious risk of bias, very serious inconsistency, and very serious imprecision) |
| | 1 study (400 participants) | 0.63 (95% CI 0.49 to 0.77) Men 74 or younger<br><br>0.67 (95% CI 0.59 to 0.75) Men 75 or older | |
| ***MORES*** | | | |
| Sensitivity | 4 studies (6,285 participants) | Sensitivity range 0.655 to 0.960 (range of 95% CI 0.520 to 0.990) | **Very low certainty that MORES identifies men at high risk of osteoporosis** (rated down for serious risk of bias, serious inconsistency, very serious imprecision) |

| Outcome | Number of Studies (N) | Findings | Certainty of Evidence (Rationale) |
|---|---|---|---|
| Specificity | 4 studies (6,285 participants) | Specificity range 0.580 to 0.700 (range of 95% CI 0.530 to 0.740) | **Very low certainty that MORES identifies men at high risk of osteoporosis** (rated down for serious risk of bias, serious indirectness, serious inconsistency, and very serious imprecision) |
| AUC | 4 studies (6,285 participants) | AUC range: 0.728 to 0.870 (range of 95% CI 0.693 to 0.910) | **Low certainty that MORES identifies men at high risk of osteoporosis** (rated down for serious risk of bias serious and serious imprecision) |
| | | **HIP FRACTURE** | |
| ***FRAX*** | | | |
| Sensitivity | 4 studies (59,795 participants) | Sensitivity range 0.427 to 0.900 (range of 95% CI NR to 0.940) | **Very low certainty that FRAX identifies men at high risk of hip fracture** (rated down for very serious inconsistency and serious imprecision) |
| Specificity | 4 studies (59,795 participants) | Specificity range 0.360 to 0.906 (range of 95% CI 0.350 to NR) | **Very low certainty that FRAX identifies men at high risk of hip fracture** (rated down for very serious inconsistency and serious imprecision) |
| AUC | 9 studies (506,888 participants) | AUC range: 0.674 to 0.930 (range of 95% CI 0.625 to 0.950) | **Low certainty that FRAX identifies men at high risk of hip fracture** (rated down for serious inconsistency and imprecision) |
| ***Garvan*** | | | |
| Sensitivity | 3 studies (484,846 participants) | Sensitivity range: 0.356 to 0.9 (range of 95% CI NR to 0.985 | **Low certainty that Garvan identifies men at high risk of hip fracture** (rated down for serious inconsistency, and imprecision) |
| Specificity | 3 studies (484,846 participants) | Specificity range: 0.35 to 0.908 (range of 95% CI 0.33 to NR) | **Low certainty that Garvan identifies men at high risk of hip fracture** (rated down for serious inconsistency, and imprecision) |
| AUC | 3 studies (484,846 participants) | AUC range: 0.71 to 0.773 (range of 95% CI 0.67 to 0.855) | **Moderate certainty that Garvan identifies men at high risk of hip fracture** (rated down for serious imprecision) |
| ***QFracture*** | | | |
| AUC | 5 studies (2,573,876 participants) | AUC range: 0.609 to 0.875 (range of 95% CI 0.660 to 0.863) | **Low certainty that QFracture identifies men at high risk of hip fracture** (rated down for serious risk of bias, inconsistency, and imprecision) |

| Outcome | Number of Studies (N) | Findings | Certainty of Evidence (Rationale) |
|---|---|---|---|
| **MAJOR OSTEOPOROTIC FRACTURE** | | | |
| ***FRAX*** | | | |
| Sensitivity | 3 studies (53,767 participants) | Sensitivity range 0.027 to 0.900 (range of 95% CI 0.0003 to 0.940) | **Very low certainty that FRAX identifies men at high risk of MOF** (rated down for serious risk of bias, inconsistency and imprecision) |
| Specificity | 3 studies (53,767 participants) | Specificity range 0.330 to 0.991 (range of 95% CI 0.340 to 0.996) | **Very low certainty that FRAX identifies men at high risk of MOF** (rated down for serious risk of bias, inconsistency and imprecision) |
| AUC | 9 studies (74,399 participants) | AUC range: 0.618 to 0.810 (range of 95% CI 0.620 to 0.850) | **Very low certainty that FRAX identifies men at high risk of MOF** (rated down for serious risk of bias, inconsistency and imprecision) |
| ***QFracture*** | | | |
| AUC | 5 studies (2,573,739 participants) | AUC range: 0.640 to 0.739 (range of 95% CI 0.61 to 0.746) | **Moderate certainty that QFracture identifies men at high risk of MOF** (rated down for serious imprecision) |

## Key Question 2 Summary

**KQ 2:** Among male Veterans not identified by a history of low-trauma fracture, is there a tool or combination of risk factors that identifies patients at highest risk of osteoporosis or major osteoporotic fracture?

To address this question, we evaluated the subset of studies conducted specifically in male Veterans not identified via history of low-trauma fracture. The studies described here represent a subset of the studies included above for KQ 1. These studies examined individual risk factors or risk assessment tools and their association with osteoporosis or osteopenia, defined by T-scores on DXA, and fracture defined by diagnosis codes. Eight studies[26,28,48,51,53,54,56,67] (n = 26,469) examined risk assessment tools among male Veterans, and of these, three[26,48,56] (n = 24,848) were conducted in populations of special interest where fracture risk is considered higher than the general population. Eleven studies[51,69-78] assessed individual risk factors for low BMD and/or fracture broadly categorized as medical conditions (*ie*, HIV infection, osteomyelitis, elevated BMI, chronic kidney disease, vitamin D deficiency, chronic pancreatitis), or exposures (*eg*, medication use), or combinations of conditions and exposures in specific Veteran populations. Overall, there was considerable conceptual heterogeneity across studies about how risk factors were used in tools, how independent risk factors were derived (*ie*, patient report vs electronic health record review), and how outcomes were defined (*eg*, diagnosis codes, T-scores). All included studies of risk factors were at high or unclear risk of bias.

Among male Veterans at average risk of osteoporosis and/or fracture, FRAX and OST were assessed most often. When comparing male Veterans to other male populations, FRAX and OST perform similarly. The OST/OSTA demonstrated poor-to-good discriminatory ability in

predicting osteoporosis (AUC 0.632 to 0.740) among general populations and among Veterans (AUC 0.670 to 0.890). Among general populations, FRAX demonstrated poor-to-excellent discriminatory ability (AUC 0.596 to 0.870) across all included studies and good discriminatory ability (AUC 0.72; 95% CI 0.67 to 0.78) in 1 Veteran study when predicting osteoporosis. The Mscore and VA-FARA are "homegrown" VA tools examined in 1 study each[28,67] While Mscore appears to better predict osteoporosis than VA-FARA, there are insufficient data to recommend one approach over another, or to recommend any of these tools other than FRAX and OST, which have been studied and validated across broader populations. In sum, among male Veterans at high risk for fracture, risk assessment tools had low-to-moderate discriminant validity.

Three studies[26,48,56] (n = 24,848) were conducted on populations of special interest for the VA and who may be at heightened risk of osteoporosis and fracture: 1) human immunodeficiency virus (HIV) and hepatitis C virus (HCV) infection (FRAX)[26]; 2) history of androgen deprivation therapy (ADT) for localized prostate cancer (FRAX)[48]; and 3) rheumatoid arthritis (OST).[56] Among male Veterans, FRAX appears to underestimate risk of fracture in HIV and HCV infection, as well as in those treated with ADT. Compared to its performance in average risk male Veteran populations, OST appears to perform suboptimally for predicting osteoporosis in male Veterans with rheumatoid arthritis. Among male Veterans, we identified limited evidence supporting individual risk factors for osteoporosis and/or fracture.

## Key Question 3 Summary

**KQ 3:** What system-level interventions improve uptake of osteoporosis screening among people without a history of low-trauma fracture?

In total, 20 studies were included examining system-level interventions to improve the uptake of osteoporosis screening among people without a history of low-trauma fracture. Because some studies had more than 1 active intervention arm, a total of 24 intervention arms are described across the 20 studies. Interventions for these studies fell into 8 different categories: 1) provider education (5 studies[92-96]), 2) provider and patient education (3 studies[93,97,98]), 3) provider-focused reminders (4 studies[99-102]), 4) clinical decision support tools (1 study[103]), 5) patient navigation (2 studies[97,104]), 6) patient risk assessment (3 studies[105-107]), 7) patient self-referral (4 studies[108-111]), and 8) patient-focused reminders (1 study[102]).

Overall, a majority of the identified systems-level interventions in the literature target providers (12 studies), and most of the literature excluded men or had limited males in the included samples. Overall, provider-focused approaches have mixed effectiveness in improving uptake of osteoporosis screening. Provider education-only interventions (*eg*, CME) show no improvements in uptake of osteoporosis screening (4 studies). Yet provider-focused reminder systems (4 studies) improve uptake of osteoporosis screening via BMD. Clinical decision support tools that combine tailored risk-based education for patients and tailored provider recommendations at the point of clinic visit show promise but have only been evaluated in 1 study. Combining provider interventions with targeted patient-focused approaches improves the impact of the combined intervention on uptake of osteoporosis screening.

Ten studies evaluated the effect of patient-focused approaches on uptake of osteoporosis screening. Overall, patient-focused approaches of patient navigation (2 studies), patient risk assessment (2 studies), patient reminders (1 study), and self-referral systems (4 studies) improve

osteoporosis screening via BMD. System-redesign approaches that allow patients to self-refer for screening may be more effective when using fixed appointments than open invitations to self-refer without a fixed appointment. Coupling patient approaches with provider-focused approaches only marginally increased effectiveness when compared to usual care.

Risk of developing osteoporosis, hip fracture, or MOF were deemed the outcomes critical to decision-making. Thus, these are the outcomes for which we conducted certainty of evidence (COE) ratings. These COE judgments reflect the degree of confidence we have in our summary findings. For each outcome of interest, we present the COE first by outcome modeled (*ie*, osteoporosis, hip fracture, MOF) per each of the major risk assessment tools, and then by sensitivity, specificity, and AUC. COE ratings are summarized below, with supporting information provided in Table 15.

- We found *very low* COE to support that provider education alone impacts uptake of osteoporosis screening and *low* COE when provider education is bolstered with patient educational approaches.
- We found *very low* COE for provider-focused reminders and clinical decisions support tools on the uptake of osteoporosis screening.
- For patient-focused interventions, we found at least *low* COE for all identified interventions (patient navigation, patient risk assessment, patient reminders patient self-referral).

**Table 15. Certainty of Evidence for Uptake of Osteoporosis Screening by Intervention Type**

| Intervention | Number of Studies (N) | Findings | Certainty of Evidence (Rationale) |
|---|---|---|---|
| **Provider-focused Interventions** | | | |
| Provider education | 4 randomized (14,827 participants) | Summary OR 0.98 (95% CI 0.39 to 2.50) | **Very low certainty of increased screening** (rated down for serious risk of bias, inconsistency, and imprecision) |
| | 1 nonrandomized (672 participants) | Intervention Pre: 6.9%<br>Intervention Post: 33.6%<br>Control Pre: 9.8%<br>Control Post: 34% | **Very low certainty of increased screening** (rated down for serious risk of bias, inconsistency, and imprecision) |
| Provider and patient education | 3 randomized (15,547 participants) | OR range 0.97 to 2.39 (95% CI 0.80 to 6.04) | **Low certainty of increased screening** (rated down for serious risk of bias and inconsistency) |
| Provider reminder | 2 randomized (4,044 participants) | OR range 1.29 to 5.47 (95% CI 0.94 to 15.57) | **Very low certainty of increased screening** (rated down for serious risk of bias, inconsistency, and imprecision) |

| Intervention | Number of Studies (N) | Findings | Certainty of Evidence (Rationale) |
|---|---|---|---|
| | 1 randomized (10,354 participants) | ß = -2.12 (SE 0.96) | **Very low certainty of increased screening** (rated down for serious risk of bias, inconsistency, indirectness, and very serious imprecision) |
| Provider reminder | 1 nonrandomized (4,660 participants) | 2.31 (95% CI 1.55 to 3.43) | **Very low certainty of increased screening** (rated down for serious risk of bias and inconsistency) |
| Clinical decision support tools | 1 Interrupted time series (18,309 participants) | Percent change: 3.4% (95% CI 2.03 to 4.68) | **Very low certainty of increased screening** (rated down for serious risk of bias and inconsistency) |
| **Patient-focused Interventions** | | | |
| Patient navigation | 1 randomized (119 participants) | OR 8.00 (95% CI 3.11 to 20.6) | **Low certainty of increased screening** (rated down for serious risk of bias and inconsistency) |
| | 1 Interrupted time series (564 participants) | Patient education + navigation vs usual care: p < .001 | **Very low certainty of increased screening** (rated down for serious inconsistency and indirectness) |
| Patient risk assessment | 2 randomized (262+ 4,685 participants) | Risk ratio 2.2 (95% CI 1.2 to 4.1) Percent difference 6% (p = 0.001) | **Low certainty of increased screening** (rated down for serious risk of bias and inconsistency) |
| | 1 randomized (34,229 participants) | Major osteoporotic fracture sub-hazard ratio : 0.986 Hip fracture sub-hazard ratio: 1.002 | **Low certainty of decreased fracture** (rated down for serious risk of bias and inconsistency) |
| Patient Self-referral | 4 randomized (19,840 participants) | OR range 1.31 to 4.87 (95% CI 0.96 to 7.14) | **Low certainty of increased screening** (rated down for serious risk of bias and imprecision) |
| Patient reminder | 1 randomized (10,354 participants) | Patient mailed reminder 21.4% (unadjusted) Usual Care: 10.8% (unadjusted) | **Low certainty of increased screening** (rated down for serious risk of bias and inconsistency) |

# PRIOR SYSTEMATIC REVIEWS

## Relevance to Key Question 1

Two prior systematic reviews have examined the use of screening in the prediction of osteoporosis[113] or osteoporotic fractures,[114] though only Shekelle et al[113] specifically examined

clinical (non-imaging) tools in men. Shekelle et al[113] reported low quality evidence that the OST has at least comparable accuracy in predicting DXA-determined osteoporosis in men when compared to calcaneal ultrasound. These authors also reported that while calcaneal ultrasound does not appear to be particularly effective at predicting DXA-determined osteoporosis in men, it may be a strong independent predictor of osteoporotic fracture. Allon et al[114] specifically examined the use of imaging tools like quantitative computed tomography (QCT), magnetic resonance imaging (MRI), and ultrasound as an alternative for DXA in predicting fragility fractures.[114] They recommended initial screening for risk of fracture with a risk questionnaire, with those patients identified as higher risk being referred for ultrasound, and then, if low bone density is observed, for DXA.[114]

Prior reviews only identified 1 non-invasive clinical risk prediction tool among men. We built on prior reviews by synthesizing the evidence on the utility of 18 different non-imaging clinical risk prediction tools as first-line screening for osteoporosis. Next, prior reviews found limited information on patient-important outcomes of fractures. We synthesized evidence on the utility of screening tools for risk of hip fractures (15 studies), MOF (16 studies), and osteoporosis via BMD (24 studies). Like prior reviews, we found mostly low certainty of evidence across all outcomes, regardless of the clinical risk prediction tool.

### Relevance to Key Question 2

Shekelle et al[113] is the only prior systematic review that has specifically examined risk factors for osteoporosis and osteoporotic fractures in men.[113] With respect to increased risk for osteoporosis in men, Shekelle's team found high quality evidence for an association with age, low body weight, physical inactivity, and weight loss, and moderate quality evidence for an association with spinal cord injury and with prolonged systemic corticosteroid therapy and androgen deprivation in the context of prostate cancer treatment. Low quality evidence was found suggesting that there was no association with diabetes mellitus type II and risk for osteoporosis in men. With respect to osteoporotic fractures, their review found moderate quality evidence for an association of alcohol use, and low-quality evidence for an association with spinal cord injury. One recent review narratively synthesized studies that had used artificial intelligence (AI) to identify groups at risk for osteoporosis or fractures.[115] These authors concluded that a large range of risk factors have been noted and called for a grouping of risk factors to aid in a more comprehensive approach to risk identification. Building upon the findings of Shekelle et al[113] and in line with Cruz et al[115] the current review identified risk factors for osteoporosis and osteoporotic fractures in men that can be grouped into 3 general categories: risk associated with behavior (*eg*, smoking, physical activity); risk associated with existing health conditions (*eg*, posttraumatic stress disorder, HIV); and risk associated with medications (*eg*, prednisone). We also identified risk factors and grouped similarly into factors associated with medical conditions (*eg,* HIV infection, elevated BMI, chronic kidney disease, chronic pancreatitis) or exposures (*eg*, medication use). In contrast to prior reviews, ours is the first systematic review key question to specifically focus on risk factors for osteoporosis among male Veterans.

### Relevance to Key Question 3

Within the past decade several systematic reviews have evaluated the effect of physician- and patient-directed interventions on initiation of BMD scanning in patients at risk for osteoporosis. Little and Eccles[116] reviewed RCTs evaluating a wide range of interventions to increase BMD

scanning or to initiate medication in individuals post-fracture who were at high risk for osteoporosis, including patient education or reminders, physician alerting, a combination of patient education and physician alerting, or physician and patient education. Results suggested positive effects of the interventions in increasing BMD scanning, with small to medium heterogeneity among the included studies. Laliberté et al[117] conducted a systematic review and meta-analysis of a range of interventions targeting PCPs and patients at risk for osteoporosis, most of which were multifaceted and included patient education, physician notification, and/or physician education. Results suggested that the interventions increased the incidence of BMD testing for both at-risk and high risk patients, with the effects more pronounced for high risk patients. These authors concluded that involvement of a wider range of health professionals beyond PCPs in interventions targeting osteoporosis screening and treatment may address both patient- and physician-related barriers to care. A recent Cochrane systematic review included an analysis of interventions to increase guideline-consistent behavior for the management of osteoporosis including BMD testing and prescription of osteoporosis medication. Results found high quality evidence for a GP alerting system combined with patient-directed intervention (patient education and reminder to see their GP) on improving these outcomes, and further evidence that GP alerting alone is probably effective such that adding the patient-directed component may not afford additional benefit.[118] Finally, Morfeld et al[119] included initiation of BMD testing or medication in their review of RCTs of patient education for osteoporosis prevention and found 2 studies that suggested that patient education improved these outcomes over usual care. Taken together, prior reviews have found that a range of patient- and physician-directed interventions may be effective at increasing BMD testing in individuals at risk for osteoporosis. Our review aligns with the findings of these studies and provides some nuances to the impacts of these types of interventions on uptake of osteoporosis screening via BMD. Overall, we found no evidence of impact of provider education as a standalone strategy. More robust provider-focused interventions that are integrated into clinical workflow such as clinician reminder systems and clinical decision support tools significantly increased the uptake of BMD screening. Combining provider interventions with targeted patient-focused approaches improves the impact of the combined intervention on uptake of osteoporosis screening over provider-only approaches. Patient-focused approaches of patient navigation, patient risk assessment and feedback, patient reminders, and self-referral systems improve osteoporosis screening via BMD.

## CLINICAL POLICY IMPLICATIONS

Current guidelines suggest screening for osteoporosis in men above the age of 70, and in those aged 50-69 years if additional risk factors are present, such as hypogonadism, smoking, and steroid use, among others.[5] In practice, clinical risk prediction tools can help identify patients at heightened risk of osteoporosis and/or fracture. However, these tools (*eg*, FRAX) have been developed primarily using female populations. Notably, in women above the age of 70 in the UK, systematic community-based screening using FRAX has been found to reduce the risk of hip fracture, but not major osteoporotic fracture. The benefit of systematic screening in men is unclear. Thus, the provider must weigh the risks and benefits of routine screening, including increased demand on DXA utilization and interpretation, low dose radiation with DXA, increased identification of low bone density and osteoporosis, increased use of osteoporosis medication and potential adverse events, and potential reduction in fracture risk.[112]

In the current review, we identified few high quality (low ROB) studies examining the utility of clinical risk prediction tools and individual risk factors for predicting osteoporosis and low-

trauma fracture in men. Of risk prediction tools, FRAX had the greatest number of studies among male populations. Consistent with the literature in post-menopausal women, FRAX appeared to better predict hip fracture versus MOF in men, with comparable predictive abilities to those reported in women.[120] A similar pattern was also observed with Q-fracture, which tends to perform similarly to FRAX in practice.[24] Overall, these data suggest good clinical utility of FRAX and Q-fracture in predicting fractures among men. However, it is important to note that studies examining risk prediction tools were heterogeneous in population (*ie*, fracture prevalence), follow-up length, and in how risk tools were implemented (*ie*, which risk factors were included). Interpretation of findings was further limited by missing data on sensitivity and specificity of risk prediction tools. As such, no definitive conclusions can be made regarding whether one fracture risk prediction tool performs better than another in men, and further investigation is warranted. A large gap in the literature likewise exists for the use of risk tools to predict male osteoporosis. MORES and OST appeared to predict osteoporosis reasonably well; however, most studies were at high risk of bias, and thus additional high quality studies are needed to determine the value of these tools in male populations.

It is worth noting that the purpose of OST is to screen for osteoporosis, not to assess risk of fracture, nor to trigger pharmacological interventions aimed at reducing fracture risk. As such, its good sensitivity across most studies (regardless of specificity), combined with its simplicity, renders it a useful screening tool for identifying male Veterans who may benefit from BMD testing for further risk stratification.[29] Yet across studies there is a lack of discernment between osteoporosis of the hip versus other sites when defining osteoporosis. This is important for men as low BMD of the hip better predicts fracture in men than low BMD of the spine, as men tend to develop more degenerative changes of the spine (*eg*, osteophytes) and vascular calcifications that may falsely raise BMD at this site. Therefore, focusing on tools that may predict hip osteoporosis as an outcome, as opposed to low BMD at other sites, may better identify men at higher of fracture.

Lastly, it is important to note that there is a wide range of ease of implementation of the 18 unique clinical risk prediction tools we identified in this review. Some have few, easily obtained variables like the OST which only includes age and weight. Others such as the QFracture have over a dozen variables that may require a mix of patient-reported and EHR-derived components to compute risk scores. When implementing a screening tool into clinical practice careful consideration needs to be paid to how the data to populate the clinical risk tool will be gathered (*eg*, patient-reported, health records) and integrated into clinical workflow to minimize provider and patient assessment burden. With the expanding use of natural language process (NLP) methods in medicine, such EHR-based approaches to populating risk prediction tools may become more feasible and allow for use of available EHR data to identify high-risk cases that would have otherwise gone undetected.

## LIMITATIONS

Our review has several strengths, including a protocol-driven design, a comprehensive search of nearly 9000 unique abstracts and 600 full-text reviews, inclusion of EPOC designs best suited to assess organizational-level interventions (KQ 3), and careful quality assessment. Both our review and the literature, however, have limitations. While we identified 67 unique studies, the total of identified studies when synthesized by tool and intervention for many outcomes was small, and most of literature we identified had design limitations that affected study quality. Further, 9 of

our synthesis suffered from high heterogeneity that was not easily explained via single sources and, for KQ 1, is likely attributable to differences in populating variables for risk assessment tools (*eg*, EHR-derived vs patient reported), thresholds, choice of reference databases, and prevalence in populations used to assess tools. For KQ 3, heterogeneity was attributable to a combination of intervention composition, populations assessed (*eg*, women only vs men with ADT), and timing of outcome assessments.

Importantly, a major limitation of the literature is limited evidence on the direct (*ie*, head-to-head) comparison of clinical risk prediction tools (KQ 1). Thus, we are overall less confident in the observed differences and similarities between tools assessed across different populations of men. Moreover, the majority of the studies and samples for KQ 3 were comprised of women. While we may presume system-level interventions to promote uptake of osteoporosis screening via BMD may perform similarity among men and women, there may be sex-specific differences in receptivity to these screening promotion approaches. Other limitations are detailed below.

### Publication Bias

Given the small number of studies, statistical methods to detect publication bias are not useful. Other strategies, such as searching ClinicalTrials.gov for completed but unpublished studies, are not a particularly effective way to identify publication bias.[121] Hence, we did not conduct formal publication bias analysis.

### Study Quality

We were also limited by the existing literature. Most of the identified literature was assessed as unclear or high ROB. For study quality in KQ 1 and KQ 2 of clinical risk prediction studies we used the QUADAS-2. In these studies, use of random sampling or consecutive patients for patient selection bias and potential bias introduced by knowledge of the reference standard on index test interpretation contributed to judgments of higher risk. For the 11 cohort studies that assess individual risk factors using an adapted Newcastle-Ottawa ROB approach, inadequate follow-up of cohorts was the most common ROB domain judged to be problematic, along with issues of outcome assessment (KQ 2). In KQ 3, adequacy of randomization, comparability of groups at baseline, and blinding were the greatest contributors to rating of unclear or high ROB.

### Applicability of Findings to the VA Population

Of the included studies in KQ 1 and KQ 2, 79% were conducted in studies exclusively comprised of men. For KQ 2, we focused exclusively on male Veterans. Males comprise the vast majority of VA users. Thus, the results of KQ 1 and KQ 2 are very applicable to the VA population. Yet, for KQ 3, most of the studies were among women only or among study samples dominated by women, and none were conducted in the VA health care system. While it is conceivable that systems-level interventions to promote screening behaviors may influence men and women equally, this is not known for certain. However, the findings presented here for the impact of systems-level interventions likely have applicability to any large health care system seeking to implement approaches to increase uptake of osteoporosis screening.

## RESEARCH GAPS/FUTURE RESEARCH

We identified several gaps in the existing literature that warrant further consideration. To systematically identify the existence of, and reason for, these gaps, we used an existing framework. Robinson et al[122] propose the identification of gaps categorically using the PICOTS framework (population, intervention, comparator, outcome, timing, and setting) and classification by reason (insufficient or imprecise information, biased information, inconsistency and/or not the right information). Below we apply this framework to identify the gaps in the literature for KQ 1 and KQ 2 (Table 16) and KQ 3 (Table 17).

### Table 16. Evidence Gaps and Future Research for Studies of Clinical Risk Prediction Tool and Risk Factors Among Male Veterans

| Evidence Gap | Reason | Types of Studies to Consider |
|---|---|---|
| **Population** | | |
| • Studies with more racial and ethnic diversity similar to the VA population<br>• Special populations at elevated risk of fracture such as patients with HIV, ADT, thyroid cancer, PTSD, chronic pancreatitis, and medication use (*eg,* antipsychotic use, opiates, gabapentinoids) | • Insufficient information<br>• Not the right information | • Prospective cohort studies<br>• Retrospective cohort studies<br>• Cross sectional studies |
| **Interventions** | | |
| • VA-specific tools like the VA-FARA and Mscore<br>• Garvan risk assessment tool | • Insufficient information | • Prospective cohort studies<br>• Retrospective cohort studies<br>• Cross sectional studies |
| **Comparators** | | |
| • Limited studies assessing clinical risk tool in the same population of men<br>• Limited studies assessing validity to tools derived from EHR vs patient-reported outcomes in men<br>• Limited studies using race-/ethnicity-specific reference data for BMD in men | • Insufficient information | • Prospective cohort studies<br>• Retrospective cohort studies<br>• Cross sectional studies |
| **Outcomes** | | |
| • Studies focusing on hip osteoporosis in men<br>• Sensitivity and specificity of clinical risk assessment tools for hip fracture and MOF in men | • Insufficient information | • Prospective cohort studies<br>• Retrospective cohort studies<br>• Cross sectional studies |
| **Setting** | | |
| • Large, comprehensive health care systems | • Insufficient information | • Prospective cohort studies<br>• Retrospective cohort studies |

### Table 17. Evidence Gaps and Future Research for Studies of Systems-level Approaches to Improve Uptake of Osteoporosis Screening

| Evidence Gap | Reason | Types of Studies to Consider |
|---|---|---|
| **Population** | | |

| Evidence Gap | Reason | Types of Studies to Consider |
|---|---|---|
| • Limited studies conducted with average risk male-only populations<br>• Limited studies conducted with male populations at elevated risk (*eg*, ADT) | • Insufficient information<br>• Not the right information | • Cluster-randomized trials<br>• Controlled before and after<br>• Nonrandomized trials<br>• Interrupted time series |
| **Interventions** | | |
| • Patient-focused reminders<br>• Provider-focused reminders<br>• Clinical decision support systems<br>• Patient navigation | • Insufficient information | • Cluster-randomized trials<br>• Controlled before and after<br>• Nonrandomized trials<br>• Interrupted time series |
| **Comparators** | | |
| • Comparisons of different self-referral systems (*eg,* fixed appointment vs open invitation to self-schedule)<br>• Provider reminders vs patient reminders<br>• Comparisons of different provider reminder systems (*eg*, actionable link vs static EHR-based reminders)<br>• Factorial designs to evaluate incremental benefit of provider-focused approaches added to patient-focused approaches | • Insufficient information | • Cluster-randomized trials<br>• Controlled before and after<br>• Nonrandomized trials<br>• Interrupted time series<br>• Factorial trials |
| **Outcomes** | | |
| • Implementation feasibility of embedding clinical risk assessment tools into clinic workflow<br>• Intervention cost<br>• Cost effectiveness<br>• Osteoporotic fracture<br>• Provider burden | • Insufficient information | • Cluster-randomized trials<br>• Controlled before and after<br>• Nonrandomized trials<br>• Interrupted time series<br>• Step-wedge designs |
| **Setting** | | |
| • Large, comprehensive health care systems | • Insufficient information | • Cluster-randomized trials<br>• Controlled before and after<br>• Nonrandomized trials<br>• Interrupted time series |

## CONCLUSIONS

The lifetime risk of an osteoporotic fracture in men over the age of 50 is between 20% and 30%.[1] Although this is less than the overall prevalence in women, men have higher rates of fracture-related mortality than women. Screening to identify those at highest risk of fracture-related morbidity is standard practice for women but there is uncertainty about universal screening for men. Overall, we found little evidence that directly compared risk assessment tools within the same population of men. Among the risk assessment tools assessed, the OST has good discriminatory ability in predicting osteoporosis by DXA and had 2 easily obtainable variables. Tools such as FRAX, QFracture, and Garvan display poor to excellent discrimination in predicting hip fracture and MOF. Yet the COE for these outcomes was low or very low except

for one outcome of interest (*ie*, Garvan hip fracture AUC rated moderate COE). For systems-level approaches to increase osteoporosis screening, we found that provider-focused strategies have mixed effectiveness in improving uptake of osteoporosis screening. Combining provider interventions with targeted patient-focused approaches improves the impact of the combined intervention on uptake of osteoporosis screening. Yet evidence for individual system approaches to increase osteoporosis screening is limited. When implementing a system of risk assessment and screening in a large health care system like the VA, careful consideration needs to be paid to impacts on provider workflow and patient assessment burden.

# REFERENCES


1. Watts NB. Osteoporosis in men. *Endocr Pract*. 2013;19(5):834-8.
2. Center JR, Nguyen TV, Schneider D, et al. Mortality after all major types of osteoporotic fracture in men and women: an observational study. *Lancet*. 1999;353(9156):878-82.
3. Cosman F, de Beur SJ, LeBoff MS, et al. Clinician's Guide to Prevention and Treatment of Osteoporosis. *Osteoporos Int*. 2014;25(10):2359-81.
4. Lewiecki EM, Gordon CM, Baim S, et al. International Society for Clinical Densitometry 2007 Adult and Pediatric Official Positions. *Bone*. 2008;43(6):1115-21.
5. Watts NB, Adler RA, Bilezikian JP, et al. Osteoporosis in men: an Endocrine Society clinical practice guideline. *J Clin Endocrinol Metab*. 2012;97(6):1802-22.
6. Curry SJ, Krist AH, Owens DK, et al. Screening for Osteoporosis to Prevent Fractures: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2018;319(24):2521-2531.
7. Jain S, Bilori B, Gupta A, et al. Are Men at High Risk for Osteoporosis Underscreened? A Quality Improvement Project. *Perm J*. 2016;20(1):60-64.
8. Kanis JA, Melton LJ, 3rd, Christiansen C, et al. The diagnosis of osteoporosis. *J Bone Miner Res*. 1994;9(8):1137-41.
9. Compston J, Cooper A, Cooper C, et al. UK clinical guideline for the prevention and treatment of osteoporosis. *Arch Osteoporos*. 2017;12(1):43.
10. LaFleur J, Rillamas-Sun E, Colon-Emeric CS, et al. Fracture Rates and Bone Density Among Postmenopausal Veteran and Non-Veteran Women From the Women's Health Initiative. *Gerontologist*. 2016;56 Suppl 1:S78-90.
11. Nahm ES, Charters K, Yoo E, et al. Osteoporosis Preventive Practice Between Veteran and Nonveteran Older Adults: Findings From Patient-Reported Data. *Orthop Nurs*. 2016;35(6):401-410.
12. Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264-9, W64.
13. Cochrane EPOC. Suggested risk of bias criteria for EPOC reviews. EPOC Resources for review authors, 2017. Available at: http://epoc.cochrane.org/resources/epoc-resources-review-authors. Accessed September 21, 2020.
14. Watts NB, Leslie WD, Foldes AJ, et al. 2013 International Society for Clinical Densitometry Position Development Conference: Task Force on Normative Databases. *J Clin Densitom*. 2013;16(4):472-81.
15. Evidence Partners Inc. DistillerAI website. Available at: https://www.evidencepartners.com/distiller-ai/. Accessed September 21, 2020.
16. Agency for Healthcare Research and Quality. Assessing the Accuracy of Machine-Assisted Abstract Screening With DistillerAI: A User Study. AHRQ Publication No. 19(20)-EHC026-EF. 2019.
17. Lewiecki EM, Bouchonville MF, 2nd, Chafey DH, et al. Bone Health ECHO: telementoring to improve osteoporosis care. *Women's health (London, England)*. 2016;12(1):79-81.
18. Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian journal of internal medicine*. 2013;4(2):627-635.
19. Mandrekar JN. Receiver Operating Characteristic Curve in Diagnostic Test Assessment. *J Thorac Oncol*. 2010;5(9):1315-1316.

20. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529-36.
21. Evidence Partners. Methods Commentary: Risk of Bias in Cohort Studies. Available at: https://www.evidencepartners.com/resources/methodological-resources/risk-of-bias-in-cohort-studies/. Accessed September 21, 2020.
22. Veroniki AA, Jackson D, Bender R, et al. Methods to calculate uncertainty in the estimated overall effect size from a random-effects meta-analysis. *Res Synth Methods*. 2019;10(1):23-43.
23. Hultcrantz M, Rind D, Akl EA, et al. The GRADE Working Group clarifies the construct of certainty of evidence. *J Clin Epidemiol*. 2017;87:4-13.
24. Kanis JA, Harvey NC, Johansson H, et al. Overview of Fracture Prediction Tools. *J Clin Densitom*. 2017;20(3):444-450.
25. Ettinger B, Ensrud KE, Blackwell T, et al. Performance of FRAX in a cohort of community-dwelling, ambulatory older men: the Osteoporotic Fractures in Men (MrOS) study. *Osteoporos Int*. 2013;24(4):1185-93.
26. Yin MT, Shiau S, Rimland D, et al. Fracture Prediction With Modified-FRAX in Older HIV-Infected and Uninfected Men. *J Acquir Immune Defic Syndr*. 2016;72(5):513-20.
27. Cass AR, Shepherd AJ, Asirot R, et al. Comparison of the Male Osteoporosis Risk Estimation Score (MORES) With FRAX in Identifying Men at Risk for Osteoporosis. *Ann Fam Med*. 2016;14(4):365-9.
28. Williams ST, Lawrence PT, Miller KL, et al. A comparison of electronic and manual fracture risk assessment tools in screening elderly male US veterans at risk for osteoporosis. *Osteoporos Int*. 2017;28(11):3107-3111.
29. Diem SJ, Peters KW, Gourlay ML, et al. Screening for Osteoporosis in Older Men: Operating Characteristics of Proposed Strategies for Selecting Men for BMD Testing. *J Gen Intern Med*. 2017;32(11):1235-1241.
30. Gourlay ML, Ritter VS, Fine JP, et al. Comparison of fracture risk assessment tools in older men without prior hip or spine fracture: the MrOS study. *Arch Osteoporos*. 2017;12(1):91.
31. Hamdy RC, Seier E, Whalen K, et al. FRAX calculated without BMD does not correctly identify Caucasian men with densitometric evidence of osteoporosis. *Osteoporos Int*. 2018;29(4):947-952.
32. 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.
33. Friis-Holmberg T, Rubin KH, Brixen K, et al. Fracture risk prediction using phalangeal bone mineral density or FRAX((R))?-A Danish cohort study on men and women. *J Clin Densitom*. 2014;17(1):7-15.
34. Short CE, Shaw SG, Fisher MJ, et al. Comparison of peripheral forearm DXA and clinical risk factor screening using FRAX(R) to assess the risk of HIV-associated low bone mass: a cross-sectional study. *Arch Osteoporos*. 2014;9:181.
35. Hoff M, Meyer HE, Skurtveit S, et al. Validation of FRAX and the impact of self-reported falls among elderly in a general population: the HUNT study, Norway. *Osteoporos Int*. 2017;28(10):2935-2944.
36. Marques A, Lucas R, Simoes E, et al. Do we need bone mineral density to estimate osteoporotic fracture risk? A 10-year prospective multicentre validation study. *RMD Open*. 2017;3(2):e000509.

37. Nakatoh S, Takemaru Y. Application of the fracture risk assessment tool (FRAX((R))) and determination of suitable cut-off values during primary screening in specific health check-ups in Japan. *J Bone Miner Metab*. 2013;31(6):674-80.
38. Kim JW, Koh JM, Park JH, et al. Validation of FRAX without BMD: an age-related analysis of the Fifth Korean National Health and Nutrition Examination Survey (KNHANES V-1, 2010). *Bone*. 2015;75:27-31.
39. Jang EJ, Lee YK, Choi HJ, et al. Osteoporotic Fracture Risk Assessment Using Bone Mineral Density in Korean: A Community-based Cohort Study. *J Bone Metab*. 2016;23(1):34-9.
40. Dagan N, Cohen-Stavi C, Leventer-Roberts M, et al. External validation and comparison of three prediction tools for risk of osteoporotic fractures using data from population based electronic health records: retrospective cohort study. *BMJ*. 2017;356:i6755.
41. Yang S, Leslie WD, Morin SN, et al. Administrative healthcare data applied to fracture risk assessment. *Osteoporos Int*. 2019;30(3):565-571.
42. Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. *Osteoporos Int*. 2012;23(1):75-85.
43. Holloway-Kew KL, Zhang Y, Betson AG, et al. How well do the FRAX (Australia) and Garvan calculators predict incident fractures? Data from the Geelong Osteoporosis Study. *Osteoporos Int*. 2019;30(10):2129-2139.
44. Hosmer DW, Lemeshow S, Sturdivant RX. Applied logistic regression. Third edition / ed Hoboken, New Jersey: Wiley; 2013 Wiley series in probability and statistics).
45. Leslie WD, Johansson H, McCloskey EV, et al. Comparison of Methods for Improving Fracture Risk Assessment in Diabetes: The Manitoba BMD Registry. *J Bone Miner Res*. 2018;33(11):1923-1930.
46. Kanis JA, Johansson H, Oden A, et al. Guidance for the adjustment of FRAX according to the dose of glucocorticoids. *Osteoporos Int*. 2011;22(3):809-816.
47. Leslie WD, Lix LM, Johansson H, et al. Spine-hip discordance and fracture risk assessment: a physician-friendly FRAX enhancement. *Osteoporos Int*. 2011;22(3):839-47.
48. Adler RA, Hastings FW, Petkov VI. Treatment thresholds for osteoporosis in men on androgen deprivation therapy: T-score versus FRAX. *Osteoporos Int*. 2010;21(4):647-53.
49. Subramaniam S, Ima-Nirwana S, Chin KY. Performance of osteoporosis self-assessment tool (OST) in predicting osteoporosis—a review. *Int J Environ Res Public Health*. 2018;15(7).
50. Machado P, Coutinho M, da Silva JA. Selecting men for bone densitometry: performance of osteoporosis risk assessment tools in Portuguese men. *Osteoporos Int*. 2010;21(6):977-83.
51. Sinnott B, Kukreja S, Barengolts E. Utility of screening tools for the prediction of low bone mass in African American men. *Osteoporos Int*. 2006;17(5):684-92.
52. Skedros JG, Sybrowsky CL, Stoddard GJ. The osteoporosis self-assessment screening tool: a useful tool for the orthopaedic surgeon. *J Bone Joint Surg Am*. 2007;89(4):765-72.
53. Richards JS, Lazzari AA, Teves Qualler DA, et al. Validation of the osteoporosis self-assessment tool in US male veterans. *J Clin Densitom*. 2014;17(1):32-7.
54. Adler RA, Tran MT, Petkov VI. Performance of the Osteoporosis Self-assessment Screening Tool for osteoporosis in American men. *Mayo Clin Proc*. 2003;78(6):723-7.
55. Oh SM, Song BM, Nam BH, et al. Development and Validation of Osteoporosis Risk-Assessment Model for Korean Men. *Yonsei Med J*. 2016;57(1):187-96.

56. Richards JS, Peng J, Amdur RL, et al. Dual-energy X-ray absorptiometry and evaluation of the osteoporosis self-assessment tool in men with rheumatoid arthritis. *J Clin Densitom*. 2009;12(4):434-40.
57. 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.
58. Collins GS, Mallett S, Altman DG. Predicting risk of osteoporotic and hip fracture in the United Kingdom: prospective independent and external validation of QFractureScores. *BMJ*. 2011;342:d3651.
59. Shepherd AJ, Cass AR, Carlson CA, et al. Development and internal validation of the male osteoporosis risk estimation score. *Ann Fam Med*. 2007;5(6):540-6.
60. Shepherd AJ, Cass AR, Ray L. Determining risk of vertebral osteoporosis in men: validation of the male osteoporosis risk estimation score. *J Am Board Fam Med*. 2010;23(2):186-94.
61. Cass AR, Shepherd AJ. Validation of the Male Osteoporosis Risk Estimation Score (MORES) in a primary care setting. *J Am Board Fam Med*. 2013;26(4):436-44.
62. Garvan Institute of Medical Research. Bone Fracture Risk Calculator. 2021.
63. Nguyen ND, Frost SA, Center JR, et al. Development of prognostic nomograms for individualizing 5-year and 10-year fracture risks. *Osteoporos Int*. 2008;19(10):1431-1444.
64. Francesco L, Elisa B, Raffaella M, et al. Assessing Risk of Osteoporotic Fractures in Primary Care: Development and Validation of the FRA-HS Algorithm. *Calcif Tissue Int*. 2017;100(6):537-549.
65. Hayashi S, Hashimoto S, Kanzaki N, et al. Stem anteversion affects periprosthetic bone mineral density after total hip arthroplasty. *Hip Int*. 2016;26(3):260-4.
66. Ettinger B, Liu H, Blackwell T, et al. Validation of FRC, a fracture risk assessment tool, in a cohort of older men: the Osteoporotic Fractures in Men (MrOS) Study. *J Clin Densitom*. 2012;15(3):334-42.
67. Zimering MB, Shin JJ, Shah J, et al. Validation of a novel risk estimation tool for predicting low bone density in Caucasian and African American men veterans. *J Clin Densitom*. 2007;10(3):289-97.
68. Kim HY, Jang EJ, Park B, et al. Development of a Korean Fracture Risk Score (KFRS) for Predicting Osteoporotic Fracture Risk: Analysis of Data from the Korean National Health Insurance Service. *PLoS One*. 2016;11(7):e0158918.
69. Munigala S, Agarwal B, Gelrud A, et al. Chronic Pancreatitis and Fracture: A Retrospective, Population-Based Veterans Administration Study. *Pancreas*. 2016;45(3):355-61.
70. Papaleontiou M, Banerjee M, Reyes-Gastelum D, et al. Risk of Osteoporosis and Fractures in Patients with Thyroid Cancer: A Case-Control Study in U.S. Veterans. *Oncologist*. 2019.
71. Hsieh E, Shiau S, Nolan O, et al. Increased Fragility Fracture Rates in Older Men with Osteomyelitis. *Clin Infect Dis*. 2019.
72. Weaver J, Kawsky J, Corboy A. Antipsychotic use and fracture risk: An evaluation of incidence at a Veterans Affairs medical center. *Ment Health Clin*. 2019;9(1):6-11.
73. Hall RK, Sloane R, Pieper C, et al. Competing Risks of Fracture and Death in Older Adults with Chronic Kidney Disease. *J Am Geriatr Soc*. 2018;66(3):532-538.

74. Khan N, Abbas AM, Almukhtar RM, et al. Prevalence and predictors of low bone mineral density in males with ulcerative colitis. *J Clin Endocrinol Metab*. 2013;98(6):2368-75.
75. Womack JA, Goulet JL, Gibert C, et al. Physiologic frailty and fragility fracture in HIV-infected male veterans. *Clin Infect Dis*. 2013;56(10):1498-504.
76. Hain RE, Hoyt RE, Moore JL, et al. Potential association of posttraumatic stress disorder and decreased bone mineral density in repatriated prisoners of war. *Mil Med*. 2011;176(3):270-5.
77. Womack JA, Goulet JL, Gibert C, et al. Increased risk of fragility fractures among HIV infected compared to uninfected male veterans. *PLoS One*. 2011;6(2):e17217.
78. Akhter N, Sinnott B, Mahmood K, et al. Effects of vitamin D insufficiency on bone mineral density in African American men. *Osteoporos Int*. 2009;20(5):745-750.
79. Shahani L, Breaux K, Lin M, et al. Veterans Aging Cohort Study Index as a Marker of Bone Disease in HIV-Infected Patients. *AIDS Res Hum Retroviruses*. 2019;35(11-12):1143-1147.
80. Pettersson U, Nilsson M, Sundh V, et al. Physical activity is the strongest predictor of calcaneal peak bone mass in young Swedish men. *Osteoporos Int*. 2010;21(3):447-455.
81. Alswat KA. Gender Disparities in Osteoporosis. *J Clin Med Res*. 2017;9(5):382-387.
82. Schuit SC, van der Klift M, Weel AE, et al. Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam Study. *Bone*. 2004;34(1):195-202.
83. Trajanoska K, Schoufour JD, de Jonge EAL, et al. Fracture incidence and secular trends between 1989 and 2013 in a population based cohort: The Rotterdam Study. *Bone*. 2018;114:116-124.
84. Tate JP, Justice AC. Change in a prognostic index for survival in HIV infection after one year on cART by level of adherence. *48th Annual Meeting of the Infectious Diseases Society of America*. Vancouver, British Columbia, Canada,; 2010.
85. Akgun KM, Pisani MA, Fried TR, et al. Risk Factors for Medical Intensive Care Unit Admission in HIV Infected Veterans. *American Thoracic Society International Conference*. New Orleans, LA; 2010.
86. Oursler KK, Goulet JL, Crystal S, et al. Association of age and comorbidity with physical function in HIV-infected and uninfected patients: results from the Veterans Aging Cohort Study. *AIDS Patient Care STDS*. 2011;25(1):13-20.
87. Yadav D, Dhir R. HOW ACCURATE ARE ICD-9 CODES FOR ACUTE (AP) AND CHRONIC (CP) PANCREATITIS?-A LARGE VA HOSPITAL EXPERIENCE. *Pancreas*. 2006;33(4).
88. Institute of Medicine. The Health of Former Prisoners of War : Results from the Medical Examination Survey of Former POWs of World War II and the Korean Conflict. The National Academies Press; 1992.
89. Fujiyama K, Kiriyama T, Ito M, et al. Suppressive doses of thyroxine do not accelerate age-related bone loss in late postmenopausal women. *Thyroid*. 1995;5(1):13-7.
90. Wang LY, Smith AW, Palmer FL, et al. Thyrotropin suppression increases the risk of osteoporosis without decreasing recurrence in ATA low- and intermediate-risk patients with differentiated thyroid carcinoma. *Thyroid*. 2015;25(3):300-7.
91. Szafors P, Che H, Barnetche T, et al. Risk of fracture and low bone mineral density in adults with inflammatory bowel diseases. A systematic literature review with meta-analysis. *Osteoporos Int*. 2018;29(11):2389-2397.

92. Dolan BM, Yialamas MA, McMahon GT. A Randomized Educational Intervention Trial to Determine the Effect of Online Education on the Quality of Resident-Delivered Care. *J Grad Med Educ*. 2015;7(3):376-81.
93. Solomon DH, Katz JN, Finkelstein JS, et al. Osteoporosis improvement: a large-scale randomized controlled trial of patient and primary care physician education. *J Bone Miner Res*. 2007;22(11):1808-15.
94. Curtis JR, Westfall AO, Allison J, et al. Challenges in improving the quality of osteoporosis care for long-term glucocorticoid users: a prospective randomized trial. *Arch Intern Med*. 2007;167(6):591-6.
95. Solomon DH, Katz JN, La Tourette AM, et al. Multifaceted intervention to improve rheumatologists' management of glucocorticoid-induced osteoporosis: a randomized controlled trial. *Arthritis Rheum*. 2004;51(3):383-7.
96. Pazirandeh M. Does patient partnership in continuing medical education (CME) improve the outcome in osteoporosis management? *J Contin Educ Health Prof*. 2002;22(3):142-51.
97. Alibhai SMH, Breunis H, Timilshina N, et al. Improving bone health in men with prostate cancer receiving androgen deprivation therapy: Results of a randomized phase 2 trial. *Cancer*. 2018;124(6):1132-1140.
98. Solomon DH, Polinski JM, Stedman M, et al. Improving care of patients at-risk for osteoporosis: a randomized controlled trial. *J Gen Intern Med*. 2007;22(3):362-7.
99. El-Kareh RE, Gandhi TK, Poon EG, et al. Actionable reminders did not improve performance over passive reminders for overdue tests in the primary care setting. *J Am Med Inform Assoc*. 2011;18(2):160-3.
100. Levy BT, Hartz A, Woodworth G, et al. Interventions to improving osteoporosis screening: an Iowa Research Network (IRENE) study. *J Am Board Fam Med*. 2009;22(4):360-7.
101. Loo TS, Davis RB, Lipsitz LA, et al. Electronic medical record reminders and panel management to improve primary care of elderly patients. *Arch Intern Med*. 2011;171(17):1552-8.
102. Lafata JE, Kolk D, Peterson EL, et al. Improving osteoporosis screening: results from a randomized cluster trial. *J Gen Intern Med*. 2007;22(3):346-51.
103. Kastner M, Sawka AM, Hamid J, et al. A knowledge translation tool improved osteoporosis disease management in primary care: an interrupted time series analysis. *Implement Sci*. 2014;9:109.
104. Denberg TD, Myers BA, Lin CT, et al. An outreach intervention increases bone densitometry testing in older women. *J Am Geriatr Soc*. 2009;57(2):341-7.
105. Heyworth L, Kleinman K, Oddleifson S, et al. Comparison of interactive voice response, patient mailing, and mailed registry to encourage screening for osteoporosis: a randomized controlled trial. *Osteoporos Int*. 2014;25(5):1519-26.
106. Yuksel N, Majumdar SR, Biggs C, et al. Community pharmacist-initiated screening program for osteoporosis: randomized controlled trial. *Osteoporos Int*. 2010;21(3):391-8.
107. Rubin KH, Rothmann MJ, Holmberg T, 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):567-578.
108. Warriner AH, Outman RC, Kitchin E, et al. A randomized trial of a mailed intervention and self-scheduling to improve osteoporosis screening in postmenopausal women. *J Bone Miner Res*. 2012;27(12):2603-10.

109. Warriner AH, Outman RC, Feldstein AC, et al. Effect of self-referral on bone mineral density testing and osteoporosis treatment. *Med Care*. 2014;52(8):743-50.
110. Ayoub WT, Newman ED, Blosky MA, et al. Improving detection and treatment of osteoporosis: redesigning care using the electronic medical record and shared medical appointments. *Osteoporos Int*. 2009;20(1):37-42.
111. Garton MJ, Torgerson DJ, Donaldson C, et al. Recruitment methods for screening programmes: trial of a new method within a regional osteoporosis study. *BMJ*. 1992;305(6845):82-4.
112. Shepstone L, Lenaghan E, Cooper C, et al. Screening in the community to reduce fractures in older women (SCOOP): a randomised controlled trial. *Lancet*. 2018;391(10122):741-747.
113. Shekelle P, Munjas B, Liu H, et al. (Department of Veterans Affairs). Screening Men for Osteoporosis: Who & How. 2007.
114. Allon R, Levy Y, Lavi I, et al. How to Best Predict Fragility Fractures: An Update and Systematic Review. *Isr Med Assoc J*. 2018;20(12):773-779.
115. Cruz AS, Lins HC, Medeiros RVA, et al. Artificial intelligence on the identification of risk groups for osteoporosis, a general review. *Biomed Eng Online*. 2018;17(1):12.
116. Little EA, Eccles MP. A systematic review of the effectiveness of interventions to improve post-fracture investigation and management of patients at risk of osteoporosis. *Implement Sci*. 2010;5:80.
117. Laliberte MC, Perreault S, Jouini G, et al. Effectiveness of interventions to improve the detection and treatment of osteoporosis in primary care settings: a systematic review and meta-analysis. *Osteoporos Int*. 2011;22(11):2743-68.
118. Tzortziou Brown V, Underwood M, Mohamed N, et al. Professional interventions for general practitioners on the management of musculoskeletal conditions. *Cochrane Database Syst Rev*. 2016(5):Cd007495.
119. Morfeld JC, Vennedey V, Muller D, et al. Patient education in osteoporosis prevention: a systematic review focusing on methodological quality of randomised controlled trials. *Osteoporos Int*. 2017;28(6):1779-1803.
120. Hillier TA, Cauley JA, Rizzo JH, et al. WHO absolute fracture risk models (FRAX): do clinical risk factors improve fracture prediction in older women without osteoporosis? *J Bone Miner Res*. 2011;26(8):1774-82.
121. Adam GP, Springs S, Trikalinos T, et al. Does information from ClinicalTrials.gov increase transparency and reduce bias? Results from a five-report case series. *Syst Rev*. 2018;7(1):59.
122. Robinson KA, Saldanha IJ, McKoy NA. Development of a framework to identify research gaps from systematic reviews. *J Clin Epidemiol*. 2011;64(12):1325-30.
123. Rubin KH, Holmberg T, Rothmann MJ, et al. The risk-stratified osteoporosis strategy evaluation study (ROSE): a randomized prospective population-based study. Design and baseline characteristics. *Calcif Tissue Int*. 2015;96(2):167-79.
124. Carey JA-O, Yang L, Erjiang E, et al. Vertebral Fractures in Ireland: A Sub-analysis of the DXA HIP Project. (1432-0827 (Electronic)).