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 Durham, NC	Conceptualization, Investigation, Methodology, Writing – original draft; Writing – review & editing
Anastasia-Stefania Alexopoulos, MBBS, MHS	Staff Endocrinologist, Durham VA Health Care System Durham, NC Assistant Professor of Medicine, Division of Endocrinology, Duke University Medical Center Durham, NC	Conceptualization, Investigation, Methodology, Writing – original draft; Writing – review & editing
Adelaide M. Gordon, MPH	Project Coordinator, Durham Evidence Synthesis Program (ESP) Center Durham, NC Research Health Science Specialist, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System 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 Durham, NC Research Health Science Specialist, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System 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 Gainesville, FL	Investigation, Validation, Writing – review & editing
Harry A Mystakelis MD, MHS	Co-Investigator, Medstar Georgetown University Hospital Department of Pediatrics 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 Durham, NC Assistant Consulting Professor, Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine Durham, NC	
Elizabeth Van Voorhees, PhD	Psychologist, Durham VA Medical Center Durham, NC 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 Durham, NC	Conceptualization, Investigation, Writing – review & editing
Richard H. Lee, MD, MPH	Co-Investigator, Durham VA Medical Center, Duke University School of Medicine 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 Associate Professor of Medicine Division of Endocrinology, Department of Medicine, Duke University School of Medicine Durham, NC	Investigation, Methodology, Writing – review & editing
Diana Soliman, MD	Fellow, Division of Endocrinology, Diabetes and Metabolism, Duke University School of Medicine Durham, NC	Investigation; Methodology, Writing – review & editing
Scott M. Carlson, MD	Fellow, Division of Endocrinology, Diabetes and Metabolism, Duke University School of Medicine 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 Durham, NC	
Sarah Cantrell, MLIS	Associate Director for Research & Education, Duke University Medical Center Library & Archives, Duke University School of Medicine Durham, NC	Conceptualization, Methodology, Writing – review & editing
Karen M. Goldstein, MD, MSPH	Co-director, Durham ESP Center Durham, NC Core Investigator, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System Durham, NC General Internist, Durham VA Medical Center Durham, NC Associate Professor, Department of Medicine, Division of General Internal Medicine, Duke University Durham, NC	Conceptualization, Methodology, Writing – review & editing
John W Williams Jr., MD, MHS	Scientific Advisor, Durham ESP Center Durham, NC Investigator, Durham Center of Innovation to Accelerate Discovery and Practice Transformation, Durham VA Health Care System Durham, NC Staff Physician, Durham VA Medical Center Durham, NC Professor, Department of Medicine, Division of General Internal Medicine, Duke University 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 Durham, NC	
	Associate Professor, Department of Population Health Sciences, Duke University School of Medicine Durham, NC	
	Associate Professor, Department of Medicine, Division of General Internal Medicine, Duke University School of Medicine 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.

EXECUTIVE SUMMARY

INTRODUCTION

Osteoporosis is a systemic skeletal disease characterized by low bone mass and deterioration of the fine structures 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). However, larger bone size, later onset of increased bone resorption, and lower fall risk are protective factors in men leading to a lower lifetime risk of fracture: 53.2% among women versus 20.7% among men. Despite a lower risk of fracture, for unclear reasons, men have higher rates of osteoporotic fracture-related complications and mortality than women.

The United States Preventive Services Task Force (USPSTF) has found insufficient evidence to recommend routine screening of men for osteoporosis. In addition to the question of *whether* to screen men for osteoporosis, there is also uncertainty about *how* to identify men for screening when 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. Thus, screening by fracture risk first, rather than by dual-energy x-ray absorptiometry (DXA), has been proposed as an alternate means to identify men who would benefit most from interventions to prevent fracture and related adverse outcomes. Fracture risk assessment tools, such as the FRAX® tool, may identify men who are at risk for osteoporotic fracture, yet who do not have BMD-defined osteoporosis. Currently, men are often identified for osteoporosis screening and treatment because of a low-impact fracture. For primary prevention, it is critical to identify men 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, 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. It is not known whether screening for osteoporosis and/or increased fracture risk in this population will reduce the future risk of osteoporotic fracture. 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 (PROSPERO registration number CRD42020150830).

Data Sources and Searches

We conducted a primary literature search of MEDLINE® (via PubMed®), Embase (via Elsevier), and CINAHL (via EBSCO). For KQ 1 and KQ 2, we searched from inception to June 28, 2019, and for KQ 3, we searched from inception to July 22, 2019. We also examined the bibliographies of recent reviews and exemplar studies identified during the topic development process for additional relevant studies. We updated the search for KQ 1, KQ 2, and KQ 3 in MEDLINE® on February 23, 2021.

Study Selection

The major eligibility criteria for study inclusion in KQ 1 and KQ 2 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: men not identified because of a prior low-trauma fracture; and (4) outcomes: osteoporosis (*ie,* BMD T-score ≤ -2.5)**,** 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 results sections. Importantly, we did not include studies with specific eligibility for a history of low-trauma fractures (*eg,* recruiting from 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. 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. Using prespecified inclusion/exclusion criteria, investigators and the DistillerSR Artificial Intelligence tool (DistillerAI; Evidence Partners Inc., Manotick, ON, Canada) evaluated titles and abstracts to identify potentially eligible studies. Studies that met all eligibility criteria at full-text review were included for data abstraction. Disagreements at the full-text review stage were resolved via consensus or a third review acting as arbiter.

Data Abstraction and Quality Assessment

Key characteristics abstracted included patient descriptors (*eg*, age, race, Veteran status, comorbidities), risk assessment tool or risk factors, intervention characteristics for KQ 3 (*eg*,

intervention target, duration/intensity, key intervention components), comparator, and outcomes. 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.

For KQ 1 and KQ 2, we assigned a summary risk of bias score ("low risk" or "at risk") to individual studies using the QUADAS-2 for Diagnostic Accuracy Studies. For cohort and case-control studies, we used the modified Newcastle-Ottawa scales, which use an overall rating of "low risk," "unclear risk" or "high risk." We used the Cochrane EPOC risk of bias (ROB) tool for KQ 3, which applies to randomized, nonrandomized, controlled before-after, and interrupted time-series studies. The overall risk of bias rating for this tool includes "low risk of bias," "unclear risk of bias," and "high risk of bias."

Data Synthesis and Analysis

We summarized the primary literature using tables, figures, and narrative synthesis. When possible, given the volume of relevant literature *(ie*, at least 3 studies reporting the same tool and outcome), heterogeneity of the studies (*ie*, $I^2 < 90\%$), and completeness of results reporting, we conducted and reported results of meta-analyses.

When a quantitative synthesis (*ie,* meta-analysis) was not appropriate, we narratively analyzed the data. We gave more weight to evidence from higher-quality studies with more precise estimates of effect. The narrative syntheses 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 risk factors, study populations, interventions, comparators, and outcome definitions.

The certainty of evidence (COE) for each KQ was assessed using the approach described by the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) working group. In brief, this approach requires assessment of 4 domains: risk of bias, consistency, directness, and precision. Additional domains to be used when appropriate were 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 (JMG, AMG) as high, moderate, low, or very low COE.

RESULTS

Results of Literature Search

For KQ 1 and KQ 2, we identified 6,269 citations, of which 364 were reviewed at the full-text stage. Of the 48 studies retained for data abstraction, 36 were cohort studies, 11 were cross-sectional studies, and 1 was a case-control study. Nineteen studies were conducted among Veterans.

We identified 6,263 citations for KQ 3, of which 235 were reviewed at the full-text stage. Of these, 20 unique studies were retained for data abstraction. They consisted of 8 randomized

controlled trials, 6 cluster-randomized trials, 2 interrupted time-series studies, 3 nonrandomized trials, and 1 controlled before-after study. None of the studies were conducted in the VA.

Summary of Results for Key Questions

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 found 37 studies evaluating 18 different tools to assess risk for osteoporosis or fracture among men not identified because of a prior low-trauma fracture. Tools varied considerably in their complexity, ranging from the inclusion of only 2 risk factors to greater than 20. The most commonly reported tools were the OST/OSTA (n = 9), FRAX (n = 19), QFracture (n = 5), MORES (n = 4), and Garvan (n = 4). Nine studies directly compared tools within the same population.

Commonly used tools for predicting osteoporosis (*ie*, FRAX, MORES) reported AUC ranges from 0.596 to 0.870, though high levels of heterogeneity were present and most studies did not use the same reference population for T-score calculation. The OST/OSTA has good discriminatory ability in predicting osteoporosis (*ie,* T-score ≤ -2.5) by DXA based on 2 easily obtainable variables (AUC ranging from 0.632 to 0.890). Commonly used tools for predicting hip fracture and major osteoporotic fracture (MOF) (*ie*, FRAX, QFracture) 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 between studies reporting on the same tool. The FRAX risk assessment tool performed better for predicting hip fracture than MOF and osteoporosis. Limited evidence was identified for use of FRAX in special populations such as individuals with HIV or those on androgen deprivation therapy (ADT); in these populations, FRAX was generally found to perform worse among these groups than in general adult male populations. Few studies reported on modified risk assessment tools (*eg*, FRAX A, eFRAX). There were 6 tools for which we found individual evaluative studies. These 6 tools displayed variable ability to discriminate higher-risk patient populations.

*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?*

Eight studies reported on clinical risk assessment tools for osteoporosis and fracture specifically among male Veterans. Twelve studies reported on independent risk factors and their associations with osteoporosis and fracture risk specifically among male Veterans. Meta-analysis was not performed due to high heterogeneity across studies. Tools performed similarly among male Veterans compared to other male populations. The OST/OSTA predicted osteoporosis similarly among male Veterans (AUC 0.670 to 0.890) as among general male populations (AUC 0.632 to 0.740). In one study of male Veterans, FRAX had an AUC of 0.72 (95% CI 0.67 to 0.78) for predicting osteoporosis, compared to AUCs ranging from 0.596 to 0.870 in general populations. Tools used to predict osteoporosis and/or fracture in Veterans at high risk for fracture (*ie,* prior ATD therapy, people living with HIV or rheumatoid arthritis) had low/moderate discriminant validity. Overall, there are insufficient data to determine whether tools perform differently across underrepresented racial/ethnic groups, as few studies include sufficiently diverse populations, and most studies do not consider race/ethnicity-specific reference data for BMD. Among male

Veterans, we identified limited evidence of an association between individual risk factors and osteoporosis and/or fracture. All included studies of risk factors were at high or unclear risk of bias.

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: provider education (5 studies); provider and patient education (3 studies); provider-focused reminders (4 studies); clinical decision support tools (1 study); patient navigation (2 studies); patient risk assessment (3 studies); patient self-referral (4 studies); and patient-focused reminders (1 study). Due to heterogeneity in populations, interventions, and study designs, we were only able to conduct meta-analysis for 1 type of intervention. All other results were narratively synthesized.

Provider-focused approaches demonstrated mixed effectiveness in improving uptake of osteoporosis screening; however, combining provider interventions with patient education resulted in a modest but statistically significant impact. Interventions involving provider education did not increase osteoporosis screening (OR 0.98; 95% CI 0.39 to 2.50; 4 studies). Provider-focused clinical reminders (4 studies) improved uptake of osteoporosis screening, and larger effects were observed in studies that combined provider reminders with patient approaches (Range OR: 1.43 to 5.47). Clinical decision support tools that combine tailored risk-based education for patients and provider recommendations at the point of clinic visit showed promise but were only evaluated in 1 study. Overall, patient-focused approaches of patient navigation, patient risk assessment, patient reminders, and self-referral systems demonstrated modest improvements in the uptake of 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.

DISCUSSION

Key Findings

Little is known about how best to identify men at increased risk for osteoporosis or fractures. We sought to assess the effectiveness of existing screening tools; which tools and risk factors were most predictive among male Veterans; and what system-level programs were most effective at increasing screening. Overall, studies were too heterogeneous to perform quantitative synthesis. Thus, the majority of the analysis was conducted via narrative synthesis. Overall, we found few studies that directly compared risk assessment tools. Among the risk assessment tools assessed, the OST was found to identify men with osteoporosis by DXA using 2 easily obtainable variables. Other tools evaluated by more than 1 study included FRAX, MORES, Garvan, and QFracture, and generally had widely varying discrimination. Tools varied in their complexity ranging from only 2 risk factors (*ie*, OST) required for calculations to more than 20 risk factors (*ie*, QFracture). When considering tools and risk factors among male Veterans specifically, we found that FRAX and OST were the tools most commonly studied for assessment of osteoporosis and/or fracture among male Veterans and that both performed similarly, with low-to-moderate discriminatory validity across osteoporosis and fracture outcomes. We identified limited

evidence supporting individual risk factors for osteoporosis and/or fracture risk but the literature suggests that existing tools may underperform among individuals with conditions such as HIV, ADT therapy, or rheumatoid arthritis. Among the many evaluated system-level interventions to increase uptake of screening, those that combine patient and provider targets may be more effective.

Applicability and Limitations

Of the included studies in KQ 1 and KQ 2, 79% were conducted exclusively in men. For KQ 2, the focus was exclusively on male Veterans and studies were drawn from samples of VA users. The results of KQ 1 and KQ 2 apply to the population of Veterans receiving care in the VA. Yet for KQ 3, most identified studies included women only or study samples dominated by women, and none were conducted in the VA health care system. The findings presented here for the impact of system-level interventions likely have applicability to any large health care system seeking to implement approaches to increase the uptake of osteoporosis screening. However, it is unclear if they will be as effective among predominantly male populations.

Our review has several strengths, including a protocol-driven design, a comprehensive search of nearly 10,000 unique abstracts and 600 full-text reviews, the 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 number of identified studies when synthesized by tool and intervention for many outcomes was small, and most of the literature we identified had design limitations that affected study quality. Thus, we relied mostly on narrative synthesis methods. Further, many of our syntheses suffered from high heterogeneity that was not easily explained via 1 variable (*eg,* race, age) and, for KQ 1, is likely attributable to differences in populating variables for risk assessment tools (*eg*, electronic health record [EHR]-derived vs patient-reported), thresholds, choice of reference databases, gender- and race-specific reference populations, and prevalence in populations used to assess tools. Also, there were limited direct comparisons of risk assessment tools; most comparisons across instruments were indirect. 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.

Research Gaps/Future Research

Additional evidence to support optimal choice and use of clinical risk assessment tools specific to male Veterans is warranted, including head-to-head comparisons of specific tools to help guide clinical decision-making. For KQ 1 and KQ 2, future research should focus on directly comparing tools within the same population. Risk assessment tools that incorporate easily accessible EHR-derived variables could increase the feasibility and acceptability in the clinical setting. While existing studies appropriately use AUC to evaluate tool discriminatory ability and optimal and patient-relevant outcomes, such as hip fracture and MOF, future research should also include standardized thresholds for these tools and reporting of sensitivity and specificity data across tools. Overall, there are insufficient data to determine whether tools perform differently across underrepresented racial/ethnic groups, as few studies include sufficiently diverse populations, and most studies do not consider race/ethnicity-specific reference data for BMD. The VHA is particularly well-suited to conduct research on system-level interventions aimed to increase appropriate screening for male Veterans at risk for osteoporosis and fracture.

Given our findings, future work on system-level interventions may focus on combining patient- and provider-focused intervention strategies.

Conclusions

Screening to identify those at the highest risk of fracture-related morbidity is standard practice for women, but there remains uncertainty about the utility of universal screening for men. Overall, we found heterogeneous evidence comparing risk assessment tools across outcomes. The high heterogeneity is perhaps attributable to differences in variables for risk assessment tools, thresholds, choice of reference populations, and prevalence in populations used to assess tools. Few studies 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 the fewest number of variables. Tools such as FRAX, QFracture, and Garvan display a range of poor to excellent discrimination in predicting hip fracture and MOF. Yet across all outcomes we judged the COE to be low to very low, suggesting that nearly every tool requires further research. For system-level approaches to increase osteoporosis screening, we found that provider-focused approaches 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. Many patient-focused approaches (*eg*, reminders, self-referral, risk assessment, and feedback) significantly increase osteoporosis screening. Yet there is limited evidence across all these interventions. When implementing a system of risk assessment and screening in a large health care system like the VHA, careful consideration needs to be given to impacts on provider workflow and patient assessment burden when considering what clinical risk assessment tool to deploy.