Supplementary Online Content

Viswanathan M, Reddy S, Berkman N, et al. Screening to prevent osteoporotic fractures: updated evidence report and systematic review for the US Preventive Services Task Force. *JAMA*. doi:10.1001/jama.2018.6537

eMethods 1.
eMethods 2.
eTables 1-62.
eFigures 1-34.
eReferences.

This supplementary material has been provided by the authors to give readers additional information about their work.

© 2018 American Medical Association. All rights reserved.

eMethods 1. Literature Search Strategies for Primary Literature

Pubmed: Screening and Treatment

Number	Results
#1	Search ("Osteoporosis"[Mesh] OR "Fractures, Bone"[Mesh] OR "Bone Density"[Mesh])
#2	Search ("Osteoporosis"[Mesh] OR "Fractures, Bone"[Mesh] OR "Bone Density"[Mesh]) Filters: Publication date from 2009/11/01; Humans; English; Adult: 19+ years
#3	Search ("Mass Screening"[Mesh] OR "Risk Assessment"[Mesh])
#4	Search (#5 AND #6)
#5	Search (#5 AND #6) Filters: Systematic Reviews
#6	Search (("Controlled Clinical Trial" [Publication Type] OR "Randomized Controlled Trial" [Publication Type] OR "Meta-Analysis" [Publication Type] OR "Cohort Studies"[Mesh]) OR "Case-Control Studies"[Mesh] OR "Sensitivity and Specificity"[Mesh])
#7	Search (#7 AND #9)
#8	Search (#8 OR #10)
#9	Search ("Osteoporosis"[Mesh] OR "Bone Density"[Mesh] OR "Calcaneus"[Mesh])
#10	Search ("Osteoporosis"[Mesh] OR "Bone Density"[Mesh] OR "Calcaneus"[Mesh])Filters: Humans; English; Adult: 19+ years
#11	Search ("Osteoporosis"[Mesh] OR "Bone Density"[Mesh] OR "Calcaneus"[Mesh])Filters: Publication date from 2009/11/01; Humans; English; Adult: 19+ years
#12	Search ((("Ultrasonography"[Mesh]) OR "Tomography, X-Ray Computed"[Mesh]) OR ("Densitometry"[Mesh] OR "Absorptiometry, Photon"[Mesh]))
#13	Search (#17 AND #19)
#14	Search (#17 AND #19) Filters: Systematic Reviews
#15	Search (#20 AND #9)
#16	Search (("Osteoporosis"[Mesh] OR "Bone Density"[Mesh]))
#17	Search (("Osteoporosis"[Mesh] OR "Bone Density"[Mesh])) Filters: Publication date from 2016/01/01; Humans; English; Adult: 19+ years
#18	Search (((((("Diphosphonates"[Mesh]) OR "Alendronate"[Mesh] OR "risedronic acid"[Supplementary Concept]) OR "Etidronic Acid"[Mesh]) OR "ibandronic acid"[Supplementary Concept]) OR "pamidronate"[Supplementary Concept]) OR "zoledronic acid"[Supplementary Concept] OR Bone Density Conservation Agents[mesh] "Calcium Carbonate"[Mesh] OR "Estrogen Receptor Modulators"[Mesh] OR "Selective Estrogen Receptor Modulators"[Mesh]
#19	Search ((("Calcitonin"[Mesh]) OR (("Hormone Replacement Therapy"[Mesh] OR "Estrogen Replacement Therapy"[Mesh]) OR "Estradiol Congeners"[Mesh])) OR ((((("Parathyroid Hormone"[Mesh]) OR "Tamoxifen"[Mesh]) OR "Teriparatide"[Mesh] OR "Raloxifene"[Mesh]) OR "Testosterone"[Mesh]) OR “RANK ligand inhibitor” OR "estropipate" [Supplementary Concept] OR "bazedoxifene" [Supplementary Concept] OR "denosumab" [Supplementary Concept])
#20	Search (#30 OR #31)
#21	Search (#28 AND #32)
#22	Search (#28 AND #32) Filters: Systematic Reviews
#23	Search (#33 AND #9)
#24	Search (#11 OR #22 OR #36)
#25	Search "Mass Screening"[Mesh] OR "Risk Assessment"[Mesh]
#26	Search (#1 AND #39)

© 2018 American Medical Association. All rights reserved.

Pubmed: Trabecular bone score

Number	Results
#102	Search "trabecular bone score "
#105	Search ("Mass Screening"[Mesh] OR "Risk Assessment"[Mesh])
#106	Search (#102 AND #105)
#107	Search (#102 AND #105) Filters: Systematic Reviews
#108	Search (#102 AND #105) Schema: all Filters: Systematic Reviews
#109	Search (("Controlled Clinical Trial" [Publication Type] OR "Randomized Controlled Trial" [Publication Type] OR "Meta-Analysis" [Publication Type] OR "Cohort Studies"[Mesh]) OR "Case-Control Studies"[Mesh] OR "Sensitivity and Specificity"[Mesh])
#110	Search (#102 AND #109)
#114	Search (#102 AND #109) Filters: Publication date from 2009/11/01; Humans; English; Adult: 19+ years

Pubmed: Denosumab

Number	Search String
#1	Search denosumab
#4	Search "Osteoporosis"[Mesh] OR "Bone Density"[Mesh]
#5	Search (#1 AND #4)
#6	Search (("Controlled Clinical Trial" [Publication Type] OR "Randomized Controlled Trial" [Publication Type] OR "Meta-Analysis" [Publication Type] OR "Cohort Studies"[Mesh]) OR "Case-Control Studies"[Mesh] OR "Sensitivity and Specificity"[Mesh])
#7	Search (#5 AND #6)
#8	Search (#5 AND #6) Filters: Humans
#9	Search (#5 AND #6) Filters: Humans; English
#10	Search (#5 AND #6) Filters: Humans; English; Adult: 19+ years

Cochrane

Osteoporosis AND (screening OR treatment)

Embase

Osteoporosis AND (screening OR treatment)

ClinicalTrials.gov

Osteoporosis AND (screening OR treatment)

Drugs@FDA.gov

Osteoporosis AND (screening OR treatment)

© 2018 American Medical Association. All rights reserved.

HSRProj

"osteoporosis"

Cochrane Clinical Trials Registry

Osteoporosis AND (screening OR treatment)

WHO ICTRP

Osteoporosis AND (screening OR treatment)

© 2018 American Medical Association. All rights reserved.

eMethods 2. United States Preventive Services Task Force Quality Rating Criteria

Randomized controlled trials and cohort studies

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

Definition of ratings based on above criteria:

Good: Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (followup ≥80%); reliable and valid measurement instruments are used and applied equally to all groups; interventions are spelled out clearly; all important outcomes are considered; and appropriate attention to confounders in analysis. In addition, intention-to-treat analysis is used for randomized controlled trials.

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

Poor: Studies are graded "poor" if any of the following fatal flaws exists: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied equally among groups (including not masking outcome assessment); and key confounders are given little or no attention. Intention-to-treat analysis is lacking for randomized controlled trials.

© 2018 American Medical Association. All rights reserved.

Source: US Preventive Services Task Force Procedure Manual. Appendix VI. Criteria for Assessing Internal Validity of Individual Studies. Available at: https://www.uspreventiveservicestaskforce.org/Page/Name/methods-and-processes

© 2018 American Medical Association. All rights reserved.

eTable 1. Eligibility Criteria for Study Selection

Crtierion	Inclusion Criteria	Exclusion Criteria
Language	Study must be published in English	Study not published in English
Type of research	Published or unpublished original research	Nonsystematic review article, letter, or editorial; articles in which results are reported elsewhere; articles with no original data
Population	KQs 1–3: General primary care men and women age ≥40 years without history of low trauma fractures; or endocrine disorders likely to be related to metabolic bone disease, such as premature ovarian failure, hypogonadism, untreated hyperthyroidism, hyperparathyroidism, adrenal insufficiency or Cushing's syndrome; or chronic use of glucocorticoids medications (>5 mg/d oral prednisone (or equivalent) for 3 months or longer) KQs 4, 5: Majority are adults with increased fracture risk	KQs 1–5: • Majority of study population has underlying medical condition likely to be related to metabolic bone disease or is already receiving treatment for osteoporosis or has experienced a low-trauma facture • Nonhuman populations • Majority of study population comprises adults younger than age 40 years KQs 4, 5: Majority are adults with no increased fracture risk
Study design	KQs 1–3: • Randomized, controlled trials • Controlled clinical trials • Systematic reviews of trials KQs 2, 3: Observational studies other than case series and case reports KQs 4: Systematic reviews and randomized controlled trials, controlled trials published since any recent, relevant review KQ 5: Systematic reviews and randomized controlled trials, controlled trials, and observational studies published since any recent, relevant review	KQ 1: Nonrandomized, controlled trials; noncontrolled clinical trials, or nonsystematic reviews of trials KQs 2, 3: Case series, case reports KQs 4, 5: Nonsystematic reviews, case series, case reports *KQ4: Case control studies*[a]
Geographical setting	KQ1, 4, 5: U.S. adult population or comparable populations (i.e., those categorized as "Very High" on the Human Development Index, as defined by the United Nations Development Programme)[b] KQ 2 and 3: Include all geographic settings	KQ 1, 4 and 5: Settings not comparable or applicable to U.S. adult population KQ2 and 3: include all geographic settings at this time
Clinical setting	KQ 1: Primary care or primary care–like settings KQs 2–5: Primary care or primary care–like settings, specialists	KQ 1: Inpatient, medical specialty (e.g., endocrinology), or nursing home settings KQs 2–5: Inpatient or nursing home settings

© 2018 American Medical Association. All rights reserved.

eTable 1. Eligibility Criteria for Study Selection (continued)

Crtierion	Inclusion Criteria	Exclusion Criteria
Intervention	KQs 1–3: Externally validated and publicly available risk assessment instruments for low bone mass, osteoporosis, or fracture risk (interventions available in the United States) *Risk assessment tools are any paper-based or electronic approach/instrument that compiles/consolidates various demographic or clinical characteristics of an individual and compares an individual's characteristics against a threshold or guideline to make a subsequent decision for testing or treatment. Examples include age, body weight criterion, Brown's clinical risk assessment, "clinical guidelines", "case identification algorithm", Elderly Falls Screening Test, Fracture absolute risk assessment, Garvan Fracture Risk Calculator, Male Osteoporosis Risk Estimation Score (MORES), NOF guidelines, Nomograms, Osteoporosis Self-Assessment Tool; Osteoporosis Self assessment Tool for Asians (OSTA); modified OSTA, ORAI, OSIRIS, QFracture algorithm, Simple Calculated Osteoporosis Risk Estimate (SCORE)*[a] *Eligible bone measurement testing includes DXA (central or peripherally measured) and quantitative ultrasound, also include dental bone tests and trabecular bone score*[a] *KQs 4, 5: Pharmacotherapy for the treatment or prevention of osteoporosis (including bisphosphonates, estrogen agonists/antagonists, hormone therapy, parathyroid hormone, and RANK Ligand Inhibitors)* *-Note: Bazedoxifine alone is not FDA approved, calcitonin is no longer used as first-line therapy*[a]	KQs 1–3: • Not externally validated or publicly available risk assessment or bone measurement testing *specifically for osteoporosis or fracture risk*[a] • Test not widely for routine clinical use in the United States KQs 2, 3: Non-FDA approved tests for screening; biomarkers of bone metabolism, quantitative CT, MRI, hip structural analysis, structural engineering models, finite element analysis KQs 4, 5: Interventions other than those described in the inclusion criteria
Comparator	KQ 1 (control interventions): no screening group KQs 2, 3 (control interventions): *Other risk assessment/testing approach, threshold, or interval* *DXA screening at hip or lumbar spine reporting T-scores based on NHANES III US White Female reference ranges*[a] KQ 4 (control interventions): Placebo KQ 5 (control interventions): Placebo or no treatment	KQ 1 (control interventions): Lack of a no screening group (active comparator) KQs 2 (control interventions): Not an active comparator *No comparator, DXA screening at peripheral sites, other non-central DXA imaging tests (e.g., Quantitative Ultrasound), T-scores based on non-NHANES or local reference ranges*[a] KQ 3: None KQs 4, 5 (control interventions): Active comparator

© 2018 American Medical Association. All rights reserved.

eTable 1. Eligibility Criteria for Study Selection (continued)

Crtierion	Inclusion Criteria	Exclusion Criteria
Outcome	All KQs: Fractures, fracture-related morbidity, fracture-related mortality, or all-cause mortality. *Fractures include "major osteoporotic fractures", which include fractures of the hip, wrist (including distal radius), humerus, and spine/vertebral (clinically presenting). Morphometric spine/vertebral fractures will also be included but recorded separately if possible.*[a] KQs 2: • Screening test characteristics *(e.g., Youden's index, sensitivity, specificity, area under the receiver operating characteristic curve or AUC, positive predictive value, negative predictive value, diagnostic odds ratio, likelihood ratio)*[a] *and reliability (test-retest measures such as Kappa)*[a] *of risk assessment (for fractures)*[a]*, bone mass measurement (for fractures or identification of osteoporosis)*[a]*, or both (for fractures)*[a] • Fracture risk prediction characteristics (overall model performance *[Brier score, R-squared] extended measures of discrimination [concordance statistic c, discrimination slope], calibration [calibration-in-the-large, calibration slope, "goodness-of-fit" test or Hosmer-Lemeshow test], reclassfication [reclassification table, reclassification calibration, net reclassification improvement, integrated discrimination improvement]), and clinical usefulness (net benefit, decision curve analysis)*[a] • Risk assessment instruments for identifying osteoporosis: AUC for ROC curves for identifying BMD ≤-2.5 KQ 3: Harms (e.g., unnecessary radiation, labeling, anxiety, false-positive results) KQ 5: Harms (e.g., cardiovascular events, hot flashes, esophageal cancer, gastrointestinal events, osteonecrosis of the jaw, atypical fractures of the femur, rashes)	Exclude if: KQ 1 *and KQ4*: • Nonvalidated fractures *(i.e., self-reported)*[a], fracture-related morbidity, or fracture-related mortality • Bone measurement testing (T-scores, z-scores) KQs 2: Outcomes *other than screening test or risk prediction characteristics*[a] *KQ 3 and KQ5: no health outcomes excluded for harms*[a]

[a] italicized text represents additional clarification to operationalize inclusion and exclusion criteria.

[b] Very high human development index countries include Norway, Australia, Switzerland, Denmark, Netherlands, Germany, Ireland, United States, Canada, New Zealand, Singapore, Hong Kong, China (SAR), Liechtenstein, Sweden, United Kingdom, Iceland, Korea (Republic of), Israel, Luxembourg, Japan, Belgium, France, Austria, Finland, Slovenia, Spain, Italy, Czech Republic, Greece, Estonia, Brunei Darussalam, Cyprus, Qatar, Andorra, Slovakia, Poland, Lithuania, Malta, Saudi Arabia, Argentina, United Arab Emirates, Chile, Portugal, Hungary, Bahrain, Latvia, Croatia, Kuwait, Montenegro (http://hdr.undp.org/en/content/table-1-human-development-index-and-its-components).

© 2018 American Medical Association. All rights reserved.

eTable 2. KQ 1 risk of bias assessment: Part 1, Participant Selection

First Author, Year	Describe Interventions and Comparators	Study Design	FOR RCTs: Was Method of Randomization Adequate?	FOR RCTs: Was Allocation Concealment Adequate?	For RCTs: Were There Baseline Imbalances that Suggest a Problem With Randomization?
Barr, 2010[1]	G1: invitation to osteoporosis screening G2: control (no invitation to screen)	RCT parallel	Yes	Probably yes	No
Shepstone, 2017[2]	G1: invitation to osteoporosis screening G2: control (no invitation to screen)	RCT parallel	Yes	Probably yes	No

Abbreviations: DXA=dual energy x-ray absorptiometry; G=group; KQ= key question; NA=not applicable; RCT=randomized controlled trial.

© 2018 American Medical Association. All rights reserved.

eTable 3. KQ 1 risk of bias assessment: Part 2, Participant Selection

First Author, Year	FOR COHORTs: Was Selection Into the Study Unrelated to Intervention or Unrelated to Outcome?	FOR COHORTs: Do Start of Follow-up and Start of Intervention Coincide for Most Participants?	FOR COHORTs: Were Adjustment Techniques Used That are Likely to Correct for the Presence of Selection Biases?	FOR CASE-CONTROLS: Were the Controls Sampled From the Population That Gave Rise to the Cases, or Using Another Method That Avoids Selection Bias?	Bias Arising From Randomization or Selection?	Comments
Barr, 2010[1]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	NR
Shepstone, 2017[2]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Some concerns	Participants healthier than nonparticipants but also more likely to have a parental history of hip fractures; however similar prevalence of parental hip fracture between screening and control group.

Abbreviations: KQ=key question; NA=not applicable; NR=not reported.

eTable 4. KQ 1 risk of bias assessment: Part 3, Confounding

First Author, Year	FOR COHORTS AND CASE CONTROLS: Is Confounding of the Effect of Intervention Unlikely in This Study?	FOR COHORTS: Were Participants Analyzed According to Their Initial Intervention Group Throughout Follow Up?	FOR COHORT STUDIES: Were Intervention Discontinuations or Switches Unlikely to be Related to Factors That are Prognostic for the Outcome?	FOR COHORT AND CASE-CONTROL STUDIES: Did the Authors Use an Appropriate Analysis Method that Adjusted for All the Critically Important Confounding Domains?
Barr, 2010[1]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Shepstone, 2017[2]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort

Abbreviations: KQ=key question; NA=not applicable.

© 2018 American Medical Association. All rights reserved.

eTable 5. KQ 1 risk of bias assessment: Part 4, Confounding

First Author, Year	FOR COHORT STUDIES: Did the authors avoid adjusting for post intervention variables?	FOR COHORT STUDIES Did the authors use an appropriate analysis method that adjusted for all the critically important confounding domains and for time-varying confounding?	Bias arising from confounding?	Comments
Barr, 2010[1]	NA-not a cohort	NA-not a cohort	No	RCT design mitigates risk of confounding from known and unknown factors.
Shepstone, 2017[2]	NA-not a cohort	NA-not a cohort	No	RCT design mitigates risk of confounding from known and unknown factors.

Abbreviations: KQ=key question; NA=not applicable; RCT=randomized controlled trial.

eTable 6. KQ 1 risk of bias assessment: Part 5, Intervention Measurement

First Author, Year	FOR COHORTS AND CASE CONTROLS: Is Intervention Status Well Defined?	FOR COHORTS AND CASE CONTROLS: Was Information on Intervention Status Recorded at the Time of Intervention?	FOR COHORTS AND CASE CONTROLS Was Information on Intervention Status Unaffected by Knowledge of the Outcome or Risk of the Outcome?	Bias Arising From Measurement of the Intervention?	Comments
Barr, 2010[1]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	RCT Design so all items NA.
Shepstone, 2017[2]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	RCT Design so all items NA.

Abbreviations: KQ=key question; NA=not applicable; RCT=randomized controlled trial.

© 2018 American Medical Association. All rights reserved.

eTable 7. KQ 1 risk of bias assessment: Part 6, Missing Data

First Author, Year	FOR RCTS and COHORTS: What was the Overall Attrition? What was the Attrition by Group? Did Attrition Vary for Different Outcomes?	FOR RCTS and COHORTS: Did the Study have High Attrition Raising Concern for Bias?	FOR RCTS and COHORTS: Are the Proportion of Participants and Reasons for Missing Data Similar Across Interventions?	FOR CASE-CONTROLS: Are the Proportion of Participants and Reasons for Missing Data Similar Across Cases and Controls?
Barr, 2010[1]	Overall: [%] unclear. Study reports an over 60% response rate but the analysis relevant for this manuscript is the per protocol analysis, and no Ns are provided. (The "ITT" analysis compares responders in the control arm to randomized in the intervention arm and therefore is not a full representation of the randomized arms and would not qualify).	Yes	No	NA
Shepstone, 2017[2]	Overall by 60 months: 10660/12483=85%. Specific Ns vary by outcome and timing of measurement	No	Yes	NA

Abbreviations: ITT=intent to treat; KQ=key question; N=number; NA=not applicable.

eTable 8. KQ 1 risk of bias assessment: Part 7, Missing Data

First Author, Year	FOR ALL STUDIES: Were Appropriate Statistical Methods Used to Account for Missing Data?	Bias Arising from Missing Outcome Data?	Comments
Barr, 2010[1]	No	Probably yes	Although this level of attrition would be considered high for trials of treatment, it's not unreasonable given the length of followup and that this was a trial of invitation to screening.
Shepstone, 2017[2]	Yes	No	NA

Abbreviations: KQ=key question.

© 2018 American Medical Association. All rights reserved.

eTable 9. KQ 1 risk of bias assessment: Part 8, Departure from Intended Intervention

First Author, Year	FOR RCTs of Treatment (NA for Screening): Were the Patients Unaware of Their Intervention Status of Participants?	FOR ALL RCTs: Were the Trial Personnel and Clinicians Unaware of the Intervention Status of Participants?	FOR ALL STUDIES Was Intervention Fidelity Adequate?	FOR ALL STUDIES: Did the Study have Enough Cross-Overs or Contamination that Would Raise Concern for Bias?	Bias Arising from Departures from Intended Interventions?
Barr, 2010[1]	No	No	No information	No information	No information
Shepstone, 2017[2]	No	No	No information	Yes	Some concerns; no masking of participants or clinicians, standards for usual care changed over the course of the trial, potentially diluting the effect of the intervention

Abbreviations: KQ=key question; NA=not applicable; RCTs=randomized controlled trials.

eTable 10. KQ 1 risk of bias assessment: Part 9, Outcome Measurement

First Author, Year	FOR ALL STUDIES: Were Benefit Outcomes (e.g., Fractures) Adequately Described, Pre-Specified, Valid, and Reliable?	FOR ALL STUDIES: Were Similar Techniques Used Among Groups to Ascertain Benefit Outcomes?	FOR ALL STUDIES: Was the Duration of Follow-Up Adequate to Assess Benefit Outcomes?	FOR ALL STUDIES: Were Harm Outcomes Adequately Described, Valid and Reliable?
Barr, 2010[1]	Probably yes	Yes	Yes	No information
Shepstone, 2017[2]	Probably yes	Yes	Yes	Yes

Abbreviations: KQ=key question; NA=not applicable.

eTable 11. KQ 1 risk of bias assessment: Part 10, Outcome Measurement

First Author, Year	FOR ALL STUDIES: Were Similar Techniques Used Among Groups to Ascertain Harm Outcomes?	FOR ALL STUDIES: Was the Duration of follow-Up Adequate to Assess Harm Outcomes?	Bias Arising From Measurement of Outcomes?
Barr, 2010[1]	No information	No information	Probably no
Shepstone, 2017[2]	Yes	Yes	Probably no. Fractures measured from medical records, so likely to have undercounted asymptomatic vertebral fractures; as a result, the study may be have been underpowered to measure these fractures, but this is a precision issue

Abbreviations: KQ=key question; NA=not applicable.

© 2018 American Medical Association. All rights reserved.

296 **eTable 12. KQ1 risk of bias assessment: Part 11, Reporting**

First Author, Year	FOR RCTS AND COHORTS: Is the Reported Effect Estimate Unlikely to be Selected, on the Basis of the Results, From Multiple Outcomes Measurements Within the Domain, Multiple Analyses, or Different Subgroups?	FOR CASE-CONTROL STUDIES: Is the Reported Effect Estimate Unlikely to be Selected, on the Basis of the Results, From Multiple Definitions of the Intervention?	Bias arising From Selection of Reported Results?
Barr, 2010[1]	No	No	No
Shepstone, 2017[2]	No	No	No

297 **Abbreviations:** KQ=key question; NA=not applicable; RCTs=randomized controlled trials.

298 **eTable 13. KQ1 risk of bias assessment: Part 12, Overall Ratings**

First Author, Year	Rating Overall (if you Rate One of the Domains as Having Bias, the Study Cannot be High Quality).	Rating Justification	Does Quality Rating of Study Vary by Outcome?	Comments
Barr, 2010[1]	Poor	The ITT analysis is not eligible because it does not fully account for all randomized; the per-protocol analysis does not account for contamination or crossovers over the long followup period; also N and loss-to-followup for per-protocol is unclear but could at least as high as 40 percent.	No	Pulled Torgeson to fully understand randomization procedures (Torgerson DJ, Thomas RE, Campbell MK, Reid DM (1997) Randomized trial of osteoporosis screening. Use of hormone replacement therapy and quality-of-life results. Arch Intern Med 157:2121–2125)
Shepstone, 2017[2]	Fair	Some concerns regarding potential contamination because of changes in guidelines over time. As a result, the difference between usual care and the intervention arm may have been reduced.Some concerns regarding potentially selection bias (participants potentially healthier but also more likely to have parents with a history of hip fractures)	No	NA

299 **Abbreviations:** ITT=intent to treat; KQ=key question; N=number; NR=not reported.

300

© 2018 American Medical Association. All rights reserved.

eTable 14. KQ2a systematic review risk of bias assessments: Part 1, Study Eligibility Criteria

First Author, Year	Describe interventions and comparators (MUST describe usual care and combinations)	Did the review adhere to pre-defined objectives and eligibility criteria?	Were the eligibility criteria appropriate for the review question?	Were eligibility criteria unambiguous?
Crandall, 2015 [3]	Not applicable	Yes	Yes	Yes
Marques et al., 2015[4]	Fracture Risk Prediction Models	Yes	Yes	Yes
Nayak et al., 2014[5]	Osteoporosis absolute fracture risk assessment instruments	Probably yes	Yes	Yes
Rubin et al., 2013[6]	Risk assessment tools	Yes	Yes	Yes
Steurer et al., 2011[7]	Development of instruments and validation	Yes	Yes	Yes

Abbreviations: KQ=key question.

eTable 15. KQ 2 systematic review risk of bias assessments: Part 2, Study Eligibility Criteria

First Author, Year	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Concerns regarding specification of study eligibility criteria	Did the review search an appropriate range of databases/electronic sources for published and unpublished reports?
Crandall, 2015 [3]	Yes	Yes	Low	Probably no
Marques et al., 2015[4]	Yes	Yes	Low	Yes
Nayak et al., 2014[5]	Yes	Yes	Low	Yes
Rubin et al., 2013[6]	Yes	Yes	Low	Probably no
Steurer et al., 2011[7]	Yes	Yes	Low	Yes

Abbreviations: KQ=key question.

© 2018 American Medical Association. All rights reserved.

306 **eTable 16. KQ 2 systematic review risk of bias assessments: Part 3, Identification and Selection of Studies**

First Author, Year	Were methods additional to database searching used to identify relevant reports?	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Were restrictions based on date, publication format, or language appropriate?	Were efforts made to minimize error in selection of studies?
Crandall, 2015 [3]	Probably no	Yes	Yes	No information
Marques et al., 2015[4]	Yes	Yes	Yes	Yes
Nayak et al., 2014[5]	Yes	Yes	Yes	Yes
Rubin et al., 2013[6]	Yes	Yes	No	Yes
Steurer et al., 2011[7]	Yes	Yes	Yes	Yes

307 **Abbreviations:** KQ=key question.

308 **eTable 17. KQ 2 systematic review risk of bias assessments: Part 4, Data Collection and Study Appraisal**

First Author, Year	Concerns regarding methods used to identify and/or select studies	Were efforts made to minimize error in data collection?	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Were all relevant study results collected for use in the synthesis?
Crandall, 2015 [3]	Unclear or some concerns	No information	Yes	Yes
Marques et al., 2015[4]	Low	Yes	Probably yes	Yes
Nayak et al., 2014[5]	Low	Yes	Yes	Yes
Rubin et al., 2013[6]	Unclear or some concerns	No information	Yes	Yes
Steurer et al., 2011[7]	Low	Yes	Yes	Yes

309 **Abbreviations:** KQ=key question.

310

© 2018 American Medical Association. All rights reserved.

eTable 18. KQ 2 systematic review risk of bias assessments: Part 5, Identification and Selection of Studies

First Author, Year	Was risk of bias (or methodological quality) formally assessed using an appropriate tool?	Were efforts made to minimize error in risk of bias assessment?	Concerns regarding methods used to collect data and appraise studies	Did the synthesis include all studies that it should?
Crandall, 2015 [3]	No	No information	Unclear or some concerns	Yes
Marques et al., 2015[4]	Yes	Yes	Low	Yes
Nayak et al., 2014[5]	Probably yes	No information	Low	Yes
Rubin et al., 2013[6]	Yes	Yes	Low	Yes
Steurer et al., 2011[7]	Yes	No information	Low	Yes

Abbreviations: KQ=key question.

eTable 19. KQ 2 systematic review risk of bias assessments: Part 6, Synthesis and Findings

First Author, Year	Were all pre-defined analyses reported or departures explained?	Was the synthesis appropriate given the degree of similarity in the research questions, study designs and outcomes across included studies?	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Were the findings robust, e.g. as demonstrated through sensitivity analyses?
Crandall, 2015[3]	Yes	Yes	Probably yes	No information
Marques et al., 2015[4]	Probably yes	Yes	Probably no	Probably yes
Nayak et al., 2014[5]	Probably yes	Yes	Yes	No information
Rubin et al., 2013[6]	Yes	Yes	Yes	Probably yes
Steurer et al., 2011[7]	Yes	Yes	No information	No information

Abbreviations: KQ=key question.

© 2018 American Medical Association. All rights reserved.

eTable 20. KQ 2 systematic review risk of bias assessments: Part 7, Overall Ratings

First Author, Year	Were biases in primary studies minimal or addressed in the synthesis?	Concerns regarding the synthesis	Did the interpretation of findings address all of the concerns identified in Domains 1 to 4?	Was the relevance of identified studies to the review's research question appropriately considered?	Did the reviewers avoid emphasizing results on the basis of their statistical significance?	Risk of bias in the review
Crandall, 2015[3]	No	Unclear or some concerns	Probably no	Yes	Yes	Unclear or some concerns
Marques et al., 2015[4]	Probably yes	Low	Yes	Yes	Yes	Low
Nayak et al., 2014[5]	Yes	Low	Yes	Yes	Probably yes	Low
Rubin et al., 2013[6]	Yes	Low	Yes	Yes	Yes	Low
Steurer et al., 2011[7]	Yes	Unclear or some concerns	No	Yes	Yes	Unclear or some concerns

Abbreviations: KQ=key question.

© 2018 American Medical Association. All rights reserved.

eTable 21. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 1, Study Description

First Author's Last Name; Year	Patients	Index Test(s):	Reference Standard and Target Condition
Adler, 2003[8]	Men enrolled in a pulmonary clinic (January-May 2001) and a rheumatology clinic (Nov 2001-March 2002) at a single VA medical center; received questionnaire and DXA scan; patients with previous DXA testing ineligble	Osteoporosis Self-assessment Tool (OST) (risk=[(weight in kg-age in years)*0.2, truncated to integer])	DXA
Bansal, 2015 [9]	All women between the ages of 50 and 64.5 years who underwent DXA during a 6-month period (March 1, 2012–August 31, 2012) and were enrolled in a primary care practice of the Mayo Clinic in Rochester, MN	FRAX	DXA
Ben Sedrine, 2001[10]	all female patients either consulting spontaneously or referred for a BMD measurement between January 1996 and September 1999 to an outpatient osteoporosis center located at the University of Lie`ge, Lie`ge, Belgium.	SCORE	DXA
Brenneman, 2003[11]	Post-menopausal women aged 60–79 in the OPRA study	SCORE SOF-based screening tool	DXA
Cadarette, 2001[12]	Post-menopausal women in CaMOS	SCORE ABONE ORAI *weight criterion and NOF also evalated	DXA
Cadarette, 2004[13]	Women >=45 years recruited prospectively from university setting and retrospectively analyzed form family practices	ORAI OST	DXA
Cass, 2006[14]	Primary care, women	ORAI and SCORE	DXA
Cass, 2013[15]	Primary care, men	MORES	DXA
Cass, 2016[16]	Primary care, men, NHANES III	MORES, FRAX	DXA
Chan, 2006[17]	Community-based elderly women	ORAI, SCORE, ABONE, OSTA	DXA
Cook et al., 2005[18]	UK, DXA scanning clinics, patients referred from general practicioners based on 1+ clinical risk factors for OP	Two QUS systems: CUBA Clinical (BUA, VOS), Sunlight Omnisense (distal radius, proximal phalanx mid finger, mid-shaft tibia)	DXA, LS-4, and total hip
Crandall, 2014[19]	Postemenopausal women enrolled in the WHI Observational or Clinical Trial Studies	OST, SCORE, USPSTF criteria (FRAX MOF risk >=9.3%)	DXA

© 2018 American Medical Association. All rights reserved.

eTable 21. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 1, Study Description (continued)

First Author's Last Name; Year	Patients	Index Test(s):	Reference Standard and Target Condition
D'Amelio, 2005[20]	Post menopausal women referred to a university-based bone metabolic unit for DXA.	NOF, OST, ORAI	DXA T Score -2.5 or less
D'Amelio, 2013[21]	Postemenopausal women recruited from general practice	NOF ORAI OST AMMEB	DXA
Geusens, 2002 [22]	Postmenopausal women 45 years and older, US and Netherlands, 81.8% white	OST, ORAI, SCORE, SOFSURF	DXA
Gnudi, 2005[23]	Postmenopausal Italian women requiring a DXA scan	Gnudi et al. clinical prediction tool	DXA
Gourlay, 2005[24]	Post menopausal women referred for DXA scans at an outpatient osteoporosis center in Belgium, based on suspicion of osteoporosis.	OST, ORAI, SCORE	DXA T Score -2.5 or less
Gourlay, 2008 [25]	US ambulatory white women aged 65 years and older	OST, ORAI, SCORE	DXA
Harrison, 2006[26]	Caucasian females, 55-80 years (referred to clinical radiology, intended use of index test (QUS x2) underwent DXA and categorized as non-osteoporosis and osteoporosis. Subsequently underwent QUS and risk assessment using demographics and then combined algorithms-QUS used to predict osteoporosis	QUS x2	DXA
Jimenez-Nunez, 2013[27]	Women from primary and tertiary care, diagnosis, no prior testing	4 risk scores + PIXI of the heel	DXA of the hip and spine
Kung, 2003[28]	Women in Hong Kong recruited from the community	OSTA index and QUI	DXA
Kung, 2005[29]	Community of Asian (Southern Chinese) men; develop index based on clinical factors; compare clinical index with calcaneal QUS in predicting BMD (T< -2.5) by DXA	Clinical index	Calcaneal QUS; target condition - osteoporosis as determined by BMD at the hip and spine by DXA
Leslie, 2013[30]	All women aged 50-64 with medical coverage and valid DXA measurements from the lumbar spine and hip, from Manitoba, Canada	FRAX, OST	DXA
Lynn, 2008[31]	US Caucasian (4658) and Hong Kong Chinese (1914) from the MrOS study with DXA and QUS measurements to compare screening tools (OST, MOST, QUI) to DXA	OST, MOST, QUI	DXA

© 2018 American Medical Association. All rights reserved.

eTable 21. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 1, Study Description (continued)

First Author's Last Name; Year	Patients	Index Test(s):	Reference Standard and Target Condition
Machado, 2010[32]	Population-based sample of Portugese men age 50 or over	OST< 1, OSTA< 2	DXA T Score -2.5 or less at any of the three sites (LS, FN, TH) measured
Martinez-Aguila, 2007[33]	Postemenopausal women age 40 to 69 referrred to a local bone densitometry unit from local gynecologists in Spain; 24% with history of prior fracture.	ORAI (>=9), OST (<2), OSIRIS(<=1)	DXA T Score -2.5 or less at FN or LS
Mauck, 2005[34]	Population-based sample of postmenopausal women age 45 years and older in Rochester, MN	SCORE >=6 ORAI >=9 NOF >=1	DXA T Score -2.5 or less at FN or LS
McLeod, 2015[35]	Women referred for screening in Canada, no prior testing	QUS and OST	DXA
Morin, 2009[36]	Population-based sample of all women age 40 to 59 and over that received DXA testing in Manitoba, Canada. Note criteria for BMD testing in women younger than 65 include premature ovarian failure, history of steroid use, prior fracture, xray evidence of osteopenia, and other pertinent clinical risk factors.	OST <=1	DXA T Score -2.5 or less at FN or LS or Total Hip
Nguyen, 2004[37]	Women from the Dubbo Osteoporosis Epidemiology Study, a population-based cohort of men and women from Dubbo, Australia.	DOESCore, FOSTA, SOFSURF, ORAI	DXA T Score < -2.5 (Reference ranges unspecified)
Oh, 2013[38]	National, population-based health and nutrition cohort.	OSTA	DXA
Oh, 2016[39]	Population based sample of Korean men age 50 and older	OSTA	DXA
Pang, 2014[40]	Persons age 70 and over recruited from general practice, excluded persons with history of fracture	OST, FRAX w/o BMD MOF and Hip	DXA
Richards, 2014[41]	Male VA patients	OST	DXA
Richy, 2004[42]	Postmenopausal White women	OST	DXA
Shepherd, 2007[43]	Men 50 years or older with DXA scan in NHANES III	MORES	DXA
Shepherd, 2010[44]	men ≥50 included in NHANES	MORES	BMD dxa osteo

© 2018 American Medical Association. All rights reserved.

eTable 21. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 1, Study Description (continued)

First Author's Last Name; Year	Patients	Index Test(s):	Reference Standard and Target Condition
Sinnott, 2006[45]	AA men, aged 35 and older (outpatient general medicine clinics at veteran hospital; intended use of clinical assessment tools and calcaneous ultrasound compared with the reference measure of BMD by DXA; no description of presentation in article; no prior testing): index text is ultrasound of calcaneous on non-dominant foot, OUTCOME is low bone mass	Ultrasound of calcaneous on non-dominant foot	BMD by DXA at the 1) lumbar spine(L1-L4) and 2)non-dominant hip(femoral neck, trochanter, total hip)
Zimering, 2007[46]	Men age 40 years or older, ambulatory veterans attending general medicine clinics, endocrinology clinics, or osteoporosis clinics	Mscore OST MSCORE (age-weight)	DXA

Abbreviations: AA= African American; ABONE=assessing age, body size, and estrogen use; AMMEB= Age, Years after Menopause, Age at Menarche, Body Mass Index ; BMD= bone mineral density; BUA=broadband attenuation; CaMOS=Canadian Multicentre Osteoporosis Study; DOEScore =Dubbo Osteoporosis Epidemiology Score; DXA=dual energy x-ray absorptiometry; DXA T=dual energy x-ray _; FN=femoral neck; FOSTA=Female Osteoporosis Self-assessment Tool for Asia; FRAX=Fracture Risk Assessment tool; LS=lumbar spine; LS-4=lumbar spine 4; MOF= major osteoporotic fracture defined as fractures of the proximal femur, distal radius, proximal humerus, and clinical vertebral fractures; MORE=Multiple Outcomes of Raloxifene Trial; MOST=Male Osteoporosis Screening Tool; MrOS=Evaluation of osteoporosis screening tools for the osteoporotic fractures in men; MSCORE= male, simple calculated osteoporosis risk estimation, NHANES III= National Health And Nutrition Examination Survey III; NOF=National Osteoporosis Foundation; OP=osteoporosis; OPRA=Osteoporosis Population-based Risk Assessment;ORAI=Osteoporosis Risk Assessment Instrument; OST=osteoporosis self-assessment tool; QUI=ultrasound index; QUS=quantitative ultrasound; SCORE=Simple Calculated Osteoporosis Risk Estimation Tool; SOF=Study of Osteoporotic Fractures; SOFSURF=Study of Osteoporotic Fractures Simple Useful Risk Factors; UK=United Kingdom; US=United States; USPSTF=United States Preventive Services Task Force; VA=Veterans' Administration; VOS=velocity of sound; WHI=Women's Health Initiative.

© 2018 American Medical Association. All rights reserved.

eTable 22. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 2, Participant Selection

First Author's Last Name; Year	Describe Methods of Patient Selection	Was a consecutive or random sample of patients enrolled?	Was a case-control design avoided?	Did the study avoid inappropriate exclusions?
Adler, 2003[8]	Data from two cross-sectional studies conducted among patients enrolled in a pulmonary clinic (evaluated from Jan-May 2001) and a rheumatology clinic (evaluated from Nov 2001-Mar 2002) at a single VA medical center.	Unclear	Yes	Yes
Bansal, 2015[9]	Conducted retrospective record review of women ages 50–64.5 years old to determine clinical factors and FRAX scores of women undergoing a DXA at researcher's institution over a 6-month period.	Yes	Yes	Yes
Ben Sedrine, 2001[10]	Gathered data from patients consulting spontaneously or referred for a BMD measurement between Jan 1996 and Sep 1999 to outpatient osteoporosis center located at University of Liege.	Unclear	Yes	Yes
Brenneman, 2003[11]	Data from first arm of OPRA study where BME testing was aimed at all women. Eligible participants were contacted by a mailing that invited all women to receive a DXA bone scan free of charge.	Yes	Yes	Unclear
Cadarette, 2001[12]	Menopausal women aged 45 years or older with DXA data at the femoral neck were included from 6 sites in the CaMos study. In the CaMos, an age-, sex-, and region-stratefied random sample of the Canadian population was selected using a telephone-based sampling frame.	Yes	Yes	Yes

© 2018 American Medical Association. All rights reserved.

eTable 22. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 2, Participant Selection (continued)

First Author's Last Name; Year	Describe Methods of Patient Selection	Was a consecutive or random sample of patients enrolled?	Was a case-control design avoided?	Did the study avoid inappropriate exclusions?
Cadarette, 2004[13]	Two groups of women were studied. Women 45 years or older presenting for BMD testing between Nov 11, 1999 and May 25, 2000 at an ambulatory care center affiliated with the University of Toronto were recruited prospectively. Women taking bone active medications other than hormone replacement, with a prior fragility fracture, or with major risk factors for secondary osteoporosis were excluded. The records of a second group of women attending two family practice clinics affiliated with the University of Toronto were reviewed retrospectively. Women aged 45 years and older with a baseline DXA report since January 1997 were eligible.	Yes	Yes	Yes
Cass, 2006[14]	Postmenopausal women aged 45 years or older receiving usual care at a university-based family practice clinic.	Yes	Yes	Yes
Cass, 2013[15]	Cross-sectional study of men who attended primary care outpatient clinics for usual care.	Yes	Yes	Yes
Cass, 2016[16]	Men aged 50 years or older from the NHANES III data set.	Yes	Yes	Yes
Chan, 2006[17]	Chinese postmenopausal women aged 55 years and older were recruited from the Tanjong Rhu community in the eastern part of Singapore.	Unclear	Yes	Unclear
Cook et al., 2005[18]	Patients referred by general practitioner to DXA screening clinic	Unclear	Yes	Unclear

© 2018 American Medical Association. All rights reserved.

eTable 22. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 2, Participant Selection (continued)

First Author's Last Name; Year	Describe Methods of Patient Selection	Was a consecutive or random sample of patients enrolled?	Was a case-control design avoided?	Did the study avoid inappropriate exclusions?
Crandall, 2014[19]	Participants from 3 clinical centers (Tucson and Phoenix, Arizona; Pittsburgh, Pennsylvania; and Birmingham, Alabama) that were part of the WHI.	Yes	Yes	Yes
D'Amelio, 2005[20]	Postmenopausal women who came to the Department of Internal Medicine to undergo bone densitometry with DXA from Aug 10, 2003 to Sep 15, 2003.	Unclear	Yes	Yes
D'Amelio, 2013[21]	Postmenopausal women referred from 32 general practitioners. Physicians were asked to send patients according to a randomization list.	Yes	Yes	Yes
Geusens, 2002 [22]	Postmenopausal women 45 years and older from US clinics and general practice in the Netherlands	yes	yes	yes
Gnudi, 2005[23]	White, postmenopausal women living in the district of Bologna, Italy and requiring DXA for BMD measurement at both the spine and hip for clinical reasons or checkups.	Yes	Yes	Yes
Gourlay, 2005[24]	Postmenopausal women aged 45 and older either self-referred or were referred by a physician for a bone mineral density scan between january 1996 and September 1999 to an outpatient osteoporosis center at the University of Liege, Liege, Belgium.	Yes	Yes	Yes
Gourlay, 2008 [25]	US ambulatory white women aged 65 years and older, from poulatio based listings	yes	yes	yes
Harrison, 2006[26]	White Caucasian females aged 55 to 70 (mean 61, SD 4) years who were referred to Clinical Radiology, Imaging Science and Biomedical Engineering, University of Manchester for routine bone densitometry scans were invited to take part in the study	Unclear	Yes	Unclear

© 2018 American Medical Association. All rights reserved.

eTable 22. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 2, Participant Selection (continued)

First Author's Last Name; Year	Describe Methods of Patient Selection	Was a consecutive or random sample of patients enrolled?	Was a case-control design avoided?	Did the study avoid inappropriate exclusions?
Jimenez-Nunez, 2013[27]	Described as random from 2 sites	Yes	Yes	Yes
Kung, 2003[28]	Women from community, all comers who did not meet exclusion	Unclear	Yes	Yes
Kung, 2005[29]	Men from community, all comers who did not meet exclusion	Yes	Yes	Yes
Leslie, 2013[30]	From a database all DXA results performed from 1990 to March 2007 in Manitoba, Canada	Yes	Yes	Yes
Lynn, 2008[31]	US participants were recruited using population-based listings at six clinical settings in Birmingham, AL; Minneapolis, MN; Palo Alto, CA; Pittsburgh, PA; Portland, OR; and San Diego, CA. Hong Kong participants were recruited using a combination of private solicitation and public advertising from community centers, housing estates, and the general community. Men who had bilateral hip replacements or who were unable to walk without the assistance of another person were excluded.	Yes	Yes	Unclear
Machado, 2010[32]	Participants were randomly selected from a list of registered voters in Santo António dos Olivais, Coimbra, Portugal. People were invited to participate by mail explaining the nature and purposes of the study.	Yes	Yes	Yes
Martinez-Aguila, 2007[33]	Questionnaire mailed to all postmenopausal patients referred by gynecologists to the rheumatology department of the Hospital Universitari de Bellvitge.	No	Yes	Unclear

© 2018 American Medical Association. All rights reserved.

eTable 22. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 2, Participant Selection (continued)

First Author's Last Name; Year	Describe Methods of Patient Selection	Was a consecutive or random sample of patients enrolled?	Was a case-control design avoided?	Did the study avoid inappropriate exclusions?
Mauck, 2005[34]	Secondary data analysis of an existing population- based cohort of postmenopausal women in Rochester, MN who were participating in an ongoing, prospective study designed to assess osteoporosis prevalence, risk factors, and outcomes. Women were recruited from an age-stratified random sample of Rochester women using the medical records linkage system of the Rochester Epidemiology Project.	Yes	Yes	Yes
McLeod, 2015[35]	Patients referred for screening to one facility	Yes	Yes	Yes
Morin, 2009[36]	Designed retrospective historical cohort study of women aged 40 to 59 years who underwent clinical BMD testing in the province for evaluation of fracture risk using a comprehensive health care databases of the Province of Manitoba in Canada.	Yes	Yes	Unclear
Nguyen, 2004[37]	All men and women aged 60 or above living in Dubbo, Australia were invited to participate in the study.	Yes	Yes	Yes
Oh, 2013[38]	Study data is based on data acquired in the KNHANES. The KNHANES is a nationwide survey to assess the health and nutritional status of a non-institutionalized representative sample of the Korean population. A stratified, multi-stage, clustered probability sampling design was used to select participants from residential districts.	Yes	Yes	Yes

© 2018 American Medical Association. All rights reserved.

eTable 22. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 2, Participant Selection (continued)

First Author's Last Name; Year	Describe Methods of Patient Selection	Was a consecutive or random sample of patients enrolled?	Was a case-control design avoided?	Did the study avoid inappropriate exclusions?
Oh, 2016[39]	Study data is based on data acquired in the KNHANES. The KNHANES is a nationwide survey to assess the health and nutritional status of a non-institutionalized representative sample of the Korean population. A stratified, multi-stage, clustered probability sampling design was used to select participants from residential districts.	Yes	Yes	Yes
Pang, 2014[40]	The study invited the participation of GPs from outer metropolitan areas with poor access to BMD. GPs involved identified individuals aged 70 and older from their practice databases. Individuals were invited to have a BMD evaluation at no personal cost.	Yes	Yes	Yes
Park, 2003[47]	From a menopause clinic, not referred from elsewhere	Unclear	Yes	Yes
Richards, 2014[41]	Attending primary care clinics at 4 participating VA Medical Centers	Unclear	Yes	Yes
Richy, 2004[42]	Patients seen at an out-patient osteoporosis centre	Unclear	Yes	Yes
Shepherd, 2007[43]	Analysis of men aged 50 years and older included in the NHANES III data set who had a valid DXA test.	Unclear	Yes	Unclear
Shepherd, 2010[44]	Men aged 50 years and older who had been included in any of the NHANES 1999 to 2000, 2001 to 2002, and 2003 to 2004 datasets and who had a valid whole-body DXA scan.	Yes	Yes	Yes
Sinnott, 2006[45]	Participants were recruited from outpatient general medicine clinics at the Jesse Brown VA Medical Center over an 11-month period in 2004	Unclear	Yes	Yes

© 2018 American Medical Association. All rights reserved.

eTable 22. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 2, Participant Selection (continued)

First Author's Last Name; Year	Describe Methods of Patient Selection	Was a consecutive or random sample of patients enrolled?	Was a case-control design avoided?	Did the study avoid inappropriate exclusions?
Zimering, 2007[46]	Men age 40 years or older were screened by 7 investigators from a population of ambulatory, community-dwelling veterans who attended general medical clinics (70%), endocrinology clinics (20%), or osteoporosis clinics (10%) at the Department of Veterans Affairs Medical Center in Lyons, New Jersey between September 1998 and September 2000.	Unclear	Yes	Yes

Abbreviations: AL=Alabama; CA=California; DXA=dual energy x-ray absorptiometry; GP= general practitioner; KNHANES= Korea National Health and Nutrition Examination Survey; MN=Minnesota; NHANES III= National Health and Nutrition Examination Survey III; OPRA= Osteoporosis Population-based Risk Assessment; PA=Pennsylvania; SD=standard deviation; US=United States; VA=Veterans' Administration; WHI= Women's Health Initiative.

© 2018 American Medical Association. All rights reserved.

eTable 23. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 3, Participant Selection

First Author's Last Name; Year	Could the selection of patients have introduced bias?	Patient Selection Comments	Describe the Index Test and How it was Conducted and Interpreted	Were the index test results interpreted without knowledge of the results of the reference standard?
Adler, 2003[8]	Unclear	Risk of spectrum bias used this ref for patient selection methods - appears random, majority of sample (107 of 181): Adler, Osteoporosis in Pulmonary Clinic Patientsa : Does Point-of-Care Screening Predict Central Dual-Energy X-ray Absorptiometry? Chest	Yes	Unclear
Bansal, 2015[9]	Unclear	Women of this age group likely had some recognized risk of osteoporosis or fracture risk (a majority [69.7%] had a previous DXA), so potential for spectrum bias	FRAX, MOF risk >=9.3%	Unclear
Ben Sedrine, 2001[10]	Unclear	Risk of spectrum bias.	Yes	Yes
Brenneman, 2003[11]	Low	Patients recruited by mailing to random sample	Yes	Unclear
Cadarette, 2001[12]	Low	Age-, sex-, and region-stratified random sample of the Canadian population selected using telephone-based sampling frame	Yes	Unclear
Cadarette, 2004[13]	Low	NA	Yes	Unclear
Cass, 2006[14]	Low	NR	Yes	Yes
Cass, 2013[15]	Low	NR	Yes	Yes
Cass, 2016[16]	Low	NHANES III is based on a probability sample of 40,000 civilian noninstitutionalized individuals	Yes	Yes
Chan, 2006[17]	Unclear	No information on participant inclusion/exclusion criteria.	Yes	Unclear
Cook et al., 2005[18]	Unclear	Sample has potential for bias toward low BMD due to recruitment from DXA clinic (all patients referred by doctor for clinical risk factors)	Two QUS tests - CUBA clinical and Sunlight Omnisense measurements. Performed on non-dominant side with same ultrsaound gel. System quality verification tests each day.	Unclear

© 2018 American Medical Association. All rights reserved.

eTable 23. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 3, Participant Selection (continued)

First Author's Last Name; Year	Could the selection of patients have introduced bias?	Patient Selection Comments	Describe the Index Test and How it was Conducted and Interpreted	Were the index test results interpreted without knowledge of the results of the reference standard?
Crandall, 2014[19]	Low	NA	Yes	Unclear
D'Amelio, 2005[20]	Unclear	Potential for spectrum bias, given the study population was referred specifically for DXA testing, in some cases for suspected secondary osteoporosis.	Yes	Unclear
D'Amelio, 2013[21]	Low	NA	Yes	Unclear
Geusens, 2002[22]	Low	NR	OST: age and weight ORAI: age, weight, estrogen use SCORE: race, rheumatoid arthritis, history of non-traumatic fracture, HRT usage, age and weight SOFSURF: age, weight, current smoker, history of postmenopausal fracture	Unclear
Gnudi, 2005[23]	Low	Patient refered to densitometry unit, possible spectrum bias	Yes	Yes
Gourlay, 2005[24]	Unclear	Potential for spectrum bias, given the study population was referred specifically for DXA testing.	Yes	Yes
Gourlay, 2008[25]	Low	NR	OST: age and weight ORAI: age, weight, estrogen use SCORE: race, rheumatoid arthritis, history of non-traumatic fracture, HRT usage, age and weight	Low
Harrison, 2006[26]	Low	No details on setting or how participants were selected	QUS x2	Unclear
Jimenez-Nunez, 2013[27]	Low	Approach to randomization using "cards" is more casual than best practice	4 risk scores + PIXI of the heel, algorithms were developed	Yes

© 2018 American Medical Association. All rights reserved.

eTable 23. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 3, Participant Selection (continued)

First Author's Last Name; Year	Could the selection of patients have introduced bias?	Patient Selection Comments	Describe the Index Test and How it was Conducted and Interpreted	Were the index test results interpreted without knowledge of the results of the reference standard?
Kung, 2003[28]	Low	Interesting that the study claims to be early postmenopausal but the age mean is 62 which makes it seem unlikely that this is actually the case	Index characteristics through interview and qui of right heel by technician	Unclear
Kung, 2005[29]	Low	Unclear who chose to participate relative to larger group, excluded abnormal TSH group	Index developed by authors based on characteristics	Unclear
Leslie, 2013[30]	Low	NR	OST, FRAX without BMD	Unclear
Lynn, 2008[31]	Low	Only exclusions listed were hip replacement and inability to walk without a cane	OST, MOST, QUI	Unclear
Machado, 2010[32]	Low	NR	Yes	Unclear
Martinez-Aguila, 2007[33]	Unclear	Patients were all referred for DXA, so potential for spectrum bias.	Yes	Unclear
Mauck, 2005[34]	Low	NR	Yes	Unclear
McLeod, 2015[35]	Low	NA	QUS of BUA and SOS of left calcaneus & personal data based on questionnaire	Yes
Morin, 2009[36]	Unclear	Population is younger women 40-59 that received a DXA, however, in this province younger women are only eligible to have coverage for DXA testing if they have clinical risks for secondary osteoporosis, history of prior fracture, or xray evidence of osteop	Yes	Unclear
Nguyen, 2004[37]	Low	NA	Yes	Unclear
Oh, 2013[38]	Low	NA	Yes	Unclear
Oh, 2016[39]	Low	NR	Yes	Unclear
Pang, 2014[40]	Low	NA	Yes	Unclear
Park, 2003[47]	Low	NR	OSTA: age and weight.	Unclear
Richards, 2014[41]	Low	NR	OST: age and weight.	Unclear

© 2018 American Medical Association. All rights reserved.

eTable 23. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 3, Participant Selection (continued)

First Author's Last Name; Year	Could the selection of patients have introduced bias?	Patient Selection Comments	Describe the Index Test and How it was Conducted and Interpreted	Were the index test results interpreted without knowledge of the results of the reference standard?
Richy, 2004[42]	Low	NR	SCORE: race, rheumatoid arthritis, history of non-traumatic fracture, HRT usage, age and weight ORAI: age, weight, estrogen use OSIRIS: age, weight, HRT use, history of low trauma fracture OST: age and weight	Unclear
Shepherd, 2007[43]	Low	NHANES uses a complex, multistage, probability sampling design to select participants representative of the civilian, non-institutionalized population of the coterminous United States, excluding Indian reservations. (i.e. not random or consecutive sampling)	Yes	Unclear
Shepherd, 2010[44]	Low	NR	Yes	Unclear
Sinnott, 2006[45]	Low	Selection of participants may be a convenience sample but unclear. Men were recruited from general medicine clinics so selection bias likely low	Ultrasound of calcaneous on non-dominant foot	Unclear

© 2018 American Medical Association. All rights reserved.

eTable 23. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 3, Participant Selection (continued)

First Author's Last Name; Year	Could the selection of patients have introduced bias?	Patient Selection Comments	Describe the Index Test and How it was Conducted and Interpreted	Were the index test results interpreted without knowledge of the results of the reference standard?
Zimering, 2007[46]	Unclear	Convenience sample 30% came from specialty clinics (endo or OP) for total cohort, but unknown for valdiation cohort Excluded those unable to assess risk factors or DXA, though did not exclude based on known medical comorbidities or bone active medications (glucocorticoids). Reported only 14% on glucocorticoids, and 4% with RA	Yes	Unclear

Abbreviations: BMD= bone mineral density; BUA=broadband attenuation; DXA=dual energy x-ray absorptiometry; HRT=hormone replacement therapy; MOST=Male Osteoporosis Screening Tool; NA=not applicable; NHANES=National Health And Nutrition Examination Survey; NR=not reported; ORAI=Osteoporosis Risk Assessment Instrument; OSIRIS=Osteoporosis Index of Risk; OSTA=Osteoporosis Self-assessment Tool for Asians; OST=osteoporosis self-assessment tool; QUI=ultrasound index; QUS=quantitative ultrasound; RA=radiographic absorptiometry; SCORE=Simple Calculated Osteoporosis Risk Estimation Tool; SOS= speed of sound; TSH=thyroid stimulating hormone.

© 2018 American Medical Association. All rights reserved.

393 **eTable 24. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 4, Index Test**

First Author's Last Name; Year	If a threshold was used, was it pre-specified?	Could the conduct or interpretation of the index test have introduced bias?	Index Test Comments	Describe the reference standard and how it was conducted and interpreted.
Adler, 2003[8]	Yes	Low	Used three cutoffs for OST - two based on published literature, one cutoff based on what they thought was appropriate	HANES reference database for hip Hologic reference source for spine Age, gender, race of reference group not reported
Bansal, 2015[9]	Yes	Low	NA	DXA, T-score < -2.5 but no other details provided
Ben Sedrine, 2001[10]	Yes	Low	Authors did report on outcomes of clinical prediction tools using a priori cutoffs. But also did calibrate tool for this population using AUC curve.	Hologic QDR reference values specifically established for the population of Liege, Belgium (local reference values)
Brenneman, 2003[11]	Yes	Low	SCORE cutoff was recalibrated using study data to achieve sensitivity of approximately 90%. Developer cut off >=6 Study cutoff >=8	NHANES III, do not specify age or gender of reference group
Cadarette, 2001[12]	Yes	Low	Used cutoffs based on those of the developers of the study	Canadian young adult normal values at the femoral neck. (Authors note that the Canadian young adult normal reference at the femoral neck (mean [SD], 0.857 [0.125] g/cm3) is similar to that reported by NHANES III for non-Hispanic white Americans (mean [SD], 0.858 [0.120] g/cm3).
Cadarette, 2004[13]	Yes	Low	Unclear timing of DXA, reference test, in relationship to index test in prospective and retrospective parts of the study sample	Unclear
Cass, 2006[14]	Yes	Low	NR	NHANES III non-Hispanic White women age 20-29 years old.
Cass, 2013[15]	Yes	Low	NR	NHANES III non-Hispanic White women age 20-29 years old.

394
395

© 2018 American Medical Association. All rights reserved.

eTable 24. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 4, Index Test (continued)

First Author's Last Name; Year	If a threshold was used, was it pre-specified?	Could the conduct or interpretation of the index test have introduced bias?	Index Test Comments	Describe the reference standard and how it was conducted and interpreted.
Cass, 2016[16]	Yes	Low	Threshold was determined in a split sample, using a development cohort, reported in Shepherd et al, 2007.[43] This analysis focuses on the validation cohort only	NHANES III non-Hispanic White women age 20-29 years old
Chan, 2006[17]	Yes	Low	Study only reports outcomes for Femoral Neck at the prespecified thresholds, the Lumbar Spine outcomes are reported using empirically derived thresholds.	Unclear
Cook et al., 2005[18]	Yes	Unclear	Threshold question - yes and no used a 90% sensitivity threshold, but also created a cut off level based on the highest combined value of Sn and Sp.	T-scores were computed using the databases supplied with the systems
Crandall, 2014[19]	Unclear	Unclear	The study mentions the existing thresholds used for the instruments from the literature, but outcomes are not reported by these thresholds.	NHANES III normative reference database (presumably young non-hispanic white females 20-29, though this is not specifically reported)
D'Amelio, 2005[20]	Yes	Low	NR	Unclear
D'Amelio, 2013[21]	Yes	Low	The thresholds mentioned in study do not correspond entirely to thresholds used by other studies.	Unclear
Geusens, 2002 [22]	Yes	Low	NR	FN: non-hispanic female white women age 20-29 (NHANES) LS: unclear

© 2018 American Medical Association. All rights reserved.

eTable 24. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 4, Index Test (continued)

First Author's Last Name; Year	If a threshold was used, was it pre-specified?	Could the conduct or interpretation of the index test have introduced bias?	Index Test Comments	Describe the reference standard and how it was conducted and interpreted.
Gnudi, 2005[23]	Yes	Low	Do not report on blinded index test assessment. Had three apriori cutoffs from development cohort to achieve 97%, 98% and 99% sensitivity	Reference values were those reported by Norland for the European female population
Gourlay, 2005[24]	No	Unclear	Did not use pre-specified cutoffs for ORAI, OST, or SCORE. Instead, picked cut-off to achieve Sn 90% for each age group under and over 65 years. (last para p.922	T score reference range was NHANES III non-Hispanic white women age 20-29 years at the femoral neck
Gourlay, 2008[25]	Yes	Low	NR	FN: non-hispanic female white women age 20-29 (NHANES) LS: manufacturers norms for women aged 30 years
Harrison, 2006[26]	Yes	Low	NR	Hologic reference data for the T and z scores calculated using Hologic reference dataa for the lumbar spine and NHANES reference data for the proximal femur
Jimenez-Nunez, 2013[27]	Yes	Low	NR	Manufacturer's reference for the Spanish population
Kung, 2003[28]	Yes	Low	Index based on characteristics can be biased based on analysis decisions	Peak young Chinese mean values used for calculating T-scores: L1–L4 BMD
Kung, 2005[29]	Yes	Low	The authors are developing their own index test and so by definition are experimenting withtheir data	Unclear
Leslie, 2013[30]	Yes	Low	NR	Femoral T-scores calculated based on NHANES III white female reference; lumbar spine used T-scores used manufacturer's USA white female reference values

© 2018 American Medical Association. All rights reserved.

eTable 24. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 4, Index Test (continued)

First Author's Last Name; Year	If a threshold was used, was it pre-specified?	Could the conduct or interpretation of the index test have introduced bias?	Index Test Comments	Describe the reference standard and how it was conducted and interpreted.
Lynn, 2008[31]	Yes	Low	NR	US: NHANES Hong Kong: local Chinese reference ranges
Machado, 2010[32]	Yes	Low	NR	NHANES III young normal references values (sex unspecified) for FN; manufacturer's database for male Caucasian references values for LS (age unspecified)
Martinez-Aguila, 2007[33]	Yes	Low	NR	T -Scores from reference range from a study conducted in a Spanish population of healthy participants of same saex with peak bone mass
Mauck, 2005[34]	Yes	Low	NR	T scores based on references ranges for young healthy women age 20-29 years in the local community area
McLeod, 2015[35]	Yes	Low	NR	NHANES III
Morin, 2009[36]	Yes	Low	Sn and Sp reported for multiple thresholds, the threshold of <=1 is what has been used in other studies, so data was only extracted for this threshold.	Reports T Scores for LS used manufacturers US white female reference ranges, based on revised NHANES III, but these are only applicable to FN, and the study states this reference range was used for LS.
Nguyen, 2004[37]	Yes	Low	Validation cohort only.	Used BMD values of young Australian women at either the femoral neck or lumbar spine as reference to determine T score
Oh, 2013[38]	No	Unclear	The authors do not report findings for the predefined threshold of OSTA< instead they report findings for a different threshold that they selected to maximize discriminative ability.	Sexspecific norms for young Japanese women

© 2018 American Medical Association. All rights reserved.

eTable 24. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 4, Index Test (continued)

First Author's Last Name; Year	If a threshold was used, was it pre-specified?	Could the conduct or interpretation of the index test have introduced bias?	Index Test Comments	Describe the reference standard and how it was conducted and interpreted.
Oh, 2016[39]	Unclear	Unclear	Unclear whether OSTA threshold used was prespecified.	Sex specific norms for young Japanese men
Pang, 2014[40]	No	Unclear	Thresholds were not prespecified, rather they were chosen to maximize discriminative ability.	Manufacturer's sex specific normative databse and an ethnic database.
Park, 2003[47]	Yes	Low	NR	Reference range for young Korean women
Richards, 2014[41]	Yes	Low	NR	NHANES III
Richy, 2004[42]	Yes	Low	NR	Reference values specifically established for the population of Liege.
Shepherd, 2007[43]	Yes	Low	Do not report on blinded index test assessment. Threshold is determined in development cohort in this study. Applied to validation cohort.	T scores derived from race/ethnicity and sex-specific bone mineral density for Hispanic, non-Hispanic white, and non-Hispanic black men aged 20-29.
Shepherd, 2010[44]	Yes	Low	NR	White men age 20-29; whole body DXA Hologic QDR-4500A
Sinnott, 2006[45]	Unclear	Low	NR	T-scores were calculated using the manufacturer's reference values, namely a young Caucasian male database for the hip and a Caucasian female database for the spine
Zimering, 2007[46]	Yes	Low	Do not report on blinded index test assessment. Threshold is determined in development cohort in this study. Applied to validation cohort.	T score <= -2.5 compared to NHANES III young male, ethnicity/race- specific reference data

Abbreviations: AUC= area under the curve; BMD= bone mineral density; DXA=dual energy x-ray absorptiometry; GE=General Electric; NR=not reported; ORAI=Osteoporosis Risk Assessment Instrument; OST=osteoporosis self-assessment tool; OSTA=Osteoporosis Self-assessment Tool for Asians; QUI=ultrasound index; QUS=quantitative ultrasound; SCORE=Simple Calculated Osteoporosis Risk Estimation Tool; Sn=sensitivity; Sp=specificity.

© 2018 American Medical Association. All rights reserved.

eTable 25. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 5, Reference Test

First Author's Last Name; Year	Is the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Reference Standard Comments
Adler, 2003[8]	Yes	Unclear	Low	NR
Bansal, 2015[9]	Yes	Unclear	Low	NR
Ben Sedrine, 2001[10]	Yes	Yes	Low	From discussion: "All of our DXA tests come from the same densitometers and from the same clinical unit.
Brenneman, 2003[11]	Yes	Unclear	Low	NR
Cadarette, 2001[12]	Yes	Unclear	Low	NR
Cadarette, 2004[13]	Yes	Unclear	Low	Unclear timing of DXA, reference test, in relationship to index test in prospective and retrospective parts of the study sample
Cass, 2006[14]	Yes	Yes	Low	Specific reference range for T scores not reported, but used manufacturer's ranges, so likely NHANES.
Cass, 2013[15]	Yes	Yes	Low	NR
Cass, 2016[16]	Yes	Yes	Low	NR
Chan, 2006[17]	Unclear	Unclear	Unclear	No information on the specific reference ranges used to determine T-Score.
Cook et al., 2005[18]	Yes	Unclear	Unclear	NR
Crandall, 2014[19]	Yes	Unclear	Low	NR
D'Amelio, 2005[20]	Yes	Unclear	Low	No information about masking of test results, but given objective calculations that go into both the index and reference test, low chance of bias.
D'Amelio, 2013[21]	Unclear	Unclear	Unclear	Reference range for T score NR.
Geusens, 2002[22]	Yes	Unclear	Low	NR
Gnudi, 2005[23]	Yes	Yes	Low	Do not report on blinded reference test assessment.

© 2018 American Medical Association. All rights reserved.

418 **eTable 25. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 5, Reference**
419 **Test (continued)**

First Author's Last Name; Year	Is the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Reference Standard Comments
Gourlay, 2005[24]	Yes	Yes	Low	NR
Gourlay, 2008[25]	Yes	Unclear	Low	NR
Harrison, 2006[26]	Yes	Unclear	Low	NR
Jimenez-Nunez, 2013[27]	Yes	Yes	Low	NR
Kung, 2003[28]	Yes	Unclear	Low	NR
Kung, 2005[29]	Yes	Yes	Low	NR
Leslie, 2013[30]	Yes	Unclear	Low	NR
Lynn, 2008[31]	Yes	Unclear	Low	All obtained from MrOS (sequence of data collection not described)
Machado, 2010[32]	Yes	Unclear	Low	NR
Martinez-Aguila, 2007[33]	Yes	Unclear	Low	Did not use NHANES reference standards; but may be appropriate since conducted in a Spanish population.
Mauck, 2005[34]	Yes	Unclear	Low	Used a local reference range for T score values.
McLeod, 2015[35]	Yes	Yes	Low	NR
Morin, 2009[36]	Yes	Yes	Low	NR
Nguyen, 2004[37]	Yes	Unclear	Low	Local reference range for young Australian women at the FN or LS was used.
Oh, 2013[38]	Yes	Unclear	Low	NR
Oh, 2016[39]	Yes	Unclear	Low	NR
Pang, 2014[40]	Yes	Unclear	Low	NR
Park, 2003[47]	Yes	Unclear	Unclear	NR
Richards, 2014[41]	Yes	Yes	Unclear	NR
Richy, 2004[42]	Yes	Unclear	Unclear	NR
Shepherd, 2007[43]	Yes	Yes	Low	Index test was developed after DXA done, so presumably reference test interpretation blinded.
Shepherd, 2010[44]	Yes	Unclear	Low	NR

420

© 2018 American Medical Association. All rights reserved.

eTable 25. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 5, Reference Test (continued)

First Author's Last Name; Year	Is the reference standard likely to correctly classify the target condition?	Were the reference standard results interpreted without knowledge of the results of the index test?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Reference Standard Comments
Sinnott, 2006[45]	Yes	Unclear	Low	Threshold values not explicity provided.
Zimering, 2007[46]	Yes	Unclear	Low	Do not report on blinded reference test assessment.

Abbreviations: BMD= bone mineral density; DXA=dual energy x-ray absorptiometry; FN=femoral neck; LS=lumbar spine; MrOS=Evaluation of osteoporosis screening tools for the osteoporotic fractures in men; NHANES III=National Health And Nutritionexamination Survey; NR=not reported.

© 2018 American Medical Association. All rights reserved.

eTable 26. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 6, Participant Flow

First Author's Last Name; Year	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded.	Describe the time interval and any interventions between index test(s) and reference standard.	Was there an appropriate interval between index test(s) and reference standard?	Did all patients receive a reference standard?	Did patients receive the same reference standard?	Were all patients included in the analysis?
Adler, 2003[8]	Excluded patients who had previously had a DXA scan (i.e. the reference test)	1 month	Yes	Unclear	Yes	Yes
Bansal, 2015[9]	None	FRAX input collected at time of DXA or from review of medical records.	Yes	Yes	Yes	Yes
Ben Sedrine, 2001[10]	Data on those with missing data for index and DXA test were not reported	Not reported: gathered retrospective medical data on BMD measurement and risk factors between January 1996 and 1999.	Unclear	Yes	Yes	Unclear
Brenneman, 2003[11]	1986 recruited 428 consented 416 had complete data	Occurred concurrently	Yes	Yes	Yes	Yes
Cadarette, 2001[12]	69 participants missing data to calculate clinical decision rules	Not specifically reported. All baseline data collected 2/2016-9/2017, presumably includes questionnaire and DXA testing.	Unclear	Yes	Yes	No
Cadarette, 2004[13]	Of retrospective sample, 66 did not have data on estrogen use. Assumed to be negative. Only patients with DXA included.	Unclear	Unclear	Yes	Yes	No
Cass, 2006[14]	Yes	Yes	Yes	Yes	Yes	No
Cass, 2013[15]	Yes	Yes	Yes	Yes	Yes	No
Cass, 2016[16]	Details NR	NR	Unclear	Yes	Yes	Yes
Chan, 2006[17]	No	Yes	Yes	Yes	Yes	Unclear
Cook et al., 2005[18]	None	None	Yes	Yes	Yes	Yes
Crandall, 2014[19]	No	Yes	Yes	Yes	Yes	Yes
D'Amelio, 2005[20]	NR	Clinical risk factors collected at the time of DXA scan	Yes	Yes	Yes	Yes

© 2018 American Medical Association. All rights reserved.

eTable 26. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 6, Participant Flow (continued)

First Author's Last Name; Year	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded.	Describe the time interval and any interventions between index test(s) and reference standard.	Was there an appropriate interval between index test(s) and reference standard?	Did all patients receive a reference standard?	Did patients receive the same reference standard?	Were all patients included in the analysis?
D'Amelio, 2013[21]	Yes	Yes	Yes	Yes	Yes	No
Geusens, 2002 [22]	NA	unclear	unclear	yes	yes	yes
Gnudi, 2005[23]	NR	NR	Unclear	Yes	Yes	Unclear
Gourlay, 2005[24]	NR	NR	Unclear	Yes	Yes	Unclear
Gourlay, 2008 [25]	NA	unclear	unclear	yes	yes	yes
Harrison, 2006[26]	NR	NR	Unclear	Yes	Yes	Unclear
Jimenez-Nunez, 2013[27]	Nursing home, homebound, prior diagnosis of osteo, on osteo drugs, serious acute or chronic disease, hip replacement, steroids	Same day	Unclear	Yes	Yes	Unclear
Kung, 2003[28]	History or evidence of metabolic bone disease, menopause before 40, history of cancer, evidence of sig renal impariment, both hips previously fractured or replaced, prior use of any bisphosphonates, fluoride or calcitonin	NR	Unclear	Yes	Yes	Yes
Kung, 2005[29]	History or evidence of metabolic bone disease, hightory of cancer, evidence of sig renal impariment, both hips previously fractured or replaced, prior use of any bisphosphonates, fluoride or calcitonin, abnormal biochemisty including renal and liver function, serum calcium, phosphate, total alkaline phosphatase, and TSH	NR	Unclear	Yes	Yes	Yes

© 2018 American Medical Association. All rights reserved.

eTable 26. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 6, Participant Flow (continued)

First Author's Last Name; Year	Describe any patients who did not receive the index test(s) and/or reference standard or who were excluded.	Describe the time interval and any interventions between index test(s) and reference standard.	Was there an appropriate interval between index test(s) and reference standard?	Did all patients receive a reference standard?	Did patients receive the same reference standard?	Were all patients included in the analysis?
Leslie, 2013[30]	NR	NR	Unclear	Yes	Yes	Yes
Lynn, 2008[31]	NR	NR	Unclear	Yes	Yes	Na
Machado, 2010[32]	NR	NR	Unclear	Yes	Yes	Yes
Martinez-Aguila, 2007[33]	Yes	NR	Unclear	Yes	Yes	No
Mauck, 2005[34]	NR	Yes	Yes	Yes	Yes	Yes
McLeod, 2015[35]	Previous diagnosis, progressive terminal illness	Within 3 weeks	Yes	Yes	Yes	Yes
Morin, 2009[36]	NR	Unclear	Unclear	Yes	Yes	Yes
Nguyen, 2004[37]	NR	Not explicitly, but given study design presume it was concurrent.	Yes	Yes	Yes	Yes
Oh, 2013[38]	Yes	Yes	Yes	Yes	Yes	Yes
Oh, 2016[39]	Yes	Yes	Yes	Yes	Yes	Yes
Pang, 2014[40]	Yes	Yes	Yes	Yes	Yes	Yes
Park, 2003[47]	NA	Unclear	Unclear	Yes	Yes	Yes
Richards, 2014[41]	NA	Unclear	Unclear	No	Yes	No
Richy, 2004[42]	NA	Unclear	Unclear	Yes	Yes	Yes
Shepherd, 2007[43]	From Looker et al. Bone mineral measurements were performed on 3176 older men in NHANES III, but 86, or 3%, were rejected for technical reasons after review, leaving 3090 with acceptable data	NR	Unclear	Yes	Yes	Yes
Shepherd, 2010[44]	Yes	Yes	Yes	Yes	Yes	Yes
Sinnott, 2006[45]	NR	NR	Unclear	Yes	Yes	Yes
Zimering, 2007[46]	NR	Not reported, presumably concurrent testing	Unclear	Yes	Yes	No

Abbreviations: BMD= bone mineral density; DXA=dual energy x-ray absorptiometry; NHANES III=National Health And Nutrition Examination Survey III; NR=not reported; TSH=thyroid stimulating hormone.

© 2018 American Medical Association. All rights reserved.

eTable 27. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 7, Overall Ratings

First Author's Last Name; Year	Could the patient flow have introduced bias?	Patient flow Comments	Overall Judgement	Overall Comments
Adler, 2003[8]	Low	From Adler, Osteoporosis in Pulmonary Clinic Patientsa : Does Point-of-Care Screening Predict Central Dual-Energy X-ray Absorptiometry? Chest Volume 123, Issue 6, June 2003, Pages 2012–2018 98 or 107 patients received DXA scan from pulmonary cohort; unknown	Low	Unclear for domain of patient selection. Also unclear how many excluded for no DXA, but from pulmonary cohort appears small. Would give it a FAIR for ROB
Bansal, 2015[9]	Low	None	Unclear	Potential for spectrum bias because younger women with DXA likely have had some unspecified risk factors. Some risk of bias introduced by retrospective design as women age 50-64 would typically not have DXA ordered in the absence of increased risks for osteoporosis.
Ben Sedrine, 2001[10]	Unclear	No report of timing between index and reference test	Low	Risk of spectrum bias. No mention of who was excluded or if any dropped out; unclear if results looked at independently blind; Unclear for domain of flow and timing.
Brenneman, 2003[11]	Low	416 includes those with complete information not sure how many were dropped due to incomplete data; sounds like data collected all at the same time	Low	416 includes those with complete information not sure how many were dropped due to incomplete data; sounds like data collected all at the same time; not sure if blinded interpretation

© 2018 American Medical Association. All rights reserved.

eTable 27. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 7, Overall Ratings (continued)

First Author's Last Name; Year	Could the patient flow have introduced bias?	Patient flow Comments	Overall Judgement	Overall Comments
Cadarette, 2001[12]	Low	Multisite study with different DXA machines in each site. T scores were calculated from cross-calibrated Hologic BMD equivalent. Baseline period < 2 years.	Low	Unclear it assessments were blind; unclear on timing of assessments; excluded those who had osteoporosis and taking bone sparing medications, those with secondary osteoporosis, those with missing data
Cadarette, 2004[13]	Low	Study authors collected clinical risk factors taken at the same time as the DXA scan for the retrospective sample of patients For prospective study, presumably concurrent.	Low	Unclear on assessment timing; unclear on blinding; looks like those with missing data were excluded
Cass, 2006[14]	Low	23 enrolled patients did not undergo DXA scan so were not included. 173 eligible patients declined to participate.	Low	NR
Cass, 2013[15]	Low	40 patients did not undergo DXA so were dropped from the analysis.	Low	NR
Cass, 2016[16]	Low	NR	Low	NR
Chan, 2006[17]	Unclear	The Number eligible is not reported, the number of dropouts is not reported, only the final N analyzed is reported.	unclear	Some concerns in multiple domains of risk of bias lead to an overal rating of uncler.
Cook et al., 2005[18]	Low	NR	Unclear	Patient selection has the potential to skew the sample toward low BMD
Crandall, 2014[19]	Low	Main analysis was restricted to a subgroup of non HRT users by design (, supplemental analyses include HRT users and all women [including those with preventive use of HRT])	Low	NR
D'Amelio, 2005[20]	Low	NR	Low	NR

© 2018 American Medical Association. All rights reserved.

eTable 27. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 7, Overall Ratings (continued)

First Author's Last Name; Year	Could the patient flow have introduced bias?	Patient flow Comments	Overall Judgement	Overall Comments
D'Amelio, 2013[21]	Low	Some patients initially enrolled were excluded because it was determined they did not meet study criteria.	Low	NR
Geusens, 2002 [22]	unclear	Unclear because of lack of clarity around timing of the tests	unclear	No details on how the reference standard data were collected or the time interval between it and the index test
Gnudi, 2005[23]	Low	While authors don't report on timing between reference and index test, validation cohort was recruited over 6 months (<2 years)	Low	NR
Gourlay, 2005[24]	Unclear	NR	Unclear	NR
Gourlay, 2008 [25]	unclear	Unclear because of lack of clarity around timing of the tests	unclear	No details on how the reference standard data were collected or the time interval between it and the index test
Harrison, 2006[26]	Unclear	Participants underwent DXA and were categorized as non -osteo or osteo prior to QUS or risk indices	Low	Low -to-high given that osteoporosis status determined first
Jimenez-Nunez, 2013[27]	Low	random sample done with some sort of cards	Low	NR
Kung, 2003[28]	Low	NR	Low	NR
Kung, 2005[29]	Low	It is not clear what the time frame between clinical assessment of risk factors and QUS; however should be little impact;I put that all participants received the same reference standard (referring to the validated group here)	Low	NR
Leslie, 2013[30]	Low	NR	Low	NR
Lynn, 2008[31]	Low	NR	Low	Data was collected prospectively from MrOS study and then analyzed as part of this study focus.
Machado, 2010[32]	Low	Interval between clinical risks and BMD inferred to be < 2 years.	Low	NR

© 2018 American Medical Association. All rights reserved.

eTable 27. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 7, Overall Ratings (continued)

First Author's Last Name; Year	Could the patient flow have introduced bias?	Patient flow Comments	Overall Judgement	Overall Comments
Martinez-Aguila, 2007[33]	Unclear	30 eligible patients were excluded for missing data. Clinical risk factors assessed retrospectively by asking participants to answer them based on the date of their BMD testing.	Unclear	NR
Mauck, 2005[34]	Low	NR	Low	NR
McLeod, 2015[35]	Low	Effort made to contact patient, enroll and conduct OST and QUS within 3 weeks of DXA scan to complete study assessments prior to provider receiving DXA results and talking with patient.	Low	NR
Morin, 2009[36]	Unclear	Unclear for timing between DXA and index test	Unclear	NR
Nguyen, 2004[37]	Low	NR	Low	NR
Oh, 2013[38]	Low	Some patients meeting prelim criteria based on age were not eligible for a variety of reasons.	Low	Low ROB for the test thresholds used by study authors.
Oh, 2016[39]	Low	Excluded some men for probably valid reasons	Low	NR
Pang, 2014[40]	Low	Some patients meeting prelim age criteria not eligible to be included.	Low	Low ROB for the test thresholds used by study authors.
Park, 2003[47]	Unclear	Unclear because of lack of clarity around timing of the tests	Unclear	No details on how the reference standard data were collected or the time interval between it and the index tes
Richards, 2014[41]	Unclear	Unclear because of lack of clarity around timing of the tests. 2 patients were excluded from the analysis because no bmd tests were done but not the primary cause of the unclear rating	Unclear	No details on how the reference standard data were collected or the time interval between it and the index tes
Richy, 2004[42]	Unclear	Unclear because of lack of clarity around timing of the tests	Unclear	No details on how the reference standard data were collected or the time interval between it and the index tes
Shepherd, 2007[43]	Low	NR	Low	NR

© 2018 American Medical Association. All rights reserved.

eTable 27. KQ 2a Risk of Bias for Assessing Accuracy of Risk Prediction Instruments for Identifying Osteoporosis: Part 7, Overall Ratings (continued)

First Author's Last Name; Year	Could the patient flow have introduced bias?	Patient flow Comments	Overall Judgement	Overall Comments
Shepherd, 2010[44]	Low	Excluded men without DXA available, though not specifically reported NHANES enrolls participants prospectively so clinical risks and DXA likely collected concurrently.	Low	NR
Sinnott, 2006[45]	Low	The flow was not specifically described, but appears sequence was clinical assessment followed by ultrasound and then DXA.	Low	Primarily due to: 1) no information on the type of sampling. Assuming conveneience sampling; 2) not clear about the sequence of testing, but low risk of bias.
Zimering, 2007[46]	Unclear	No report of timing between index and reference test or on missing data in the validation cohort; presumably concurrent testing	Unclear	NR

Abbreviations: BMD= bone mineral density; DXA=dual energy x-ray absorptiometry; HRT=hormone replacement therapy; MrOS=Evaluation of osteoporosis screening tools for the osteoporotic fractures in men; NHANES=National Health And Nutritionexamination Survey; NR=not reported; OST=osteoporosis self-assessment tool; QUS=quantitative ultrasound; ROB=risk of bias.

© 2018 American Medical Association. All rights reserved.

460 **eTable 28. KQ 2a Risk of bias assessment forKQ 2a imaging studies predicting bone density status: Part 1, Participant Selection**

First Author, Year	Patients (Setting, Intended Use of Index Test, Presentation, Prior Testing)	Index Test(s)	Reference Standard and Target Condition	Describe Methods of Patient Selection	Was A Consecutive or Random Sample of Patients Enrolled	Was a Case-Control Design Avoided?	Did the Study Avoid Inappropriate Exclusions?	Could the Selection of Patients Have Introduced Bias?	Comments
Boonen et al., 2005[48]	Commmunty dwelling postmenopausal women,	QUS	t-score below 2.5 using dxa	Community dwelling postmenopausal women who had been referred for bone densitometry at one facility in Belgium	Yes	Yes	Yes	Low	NR
Cook et al., 2005[18]	UK, DXA scanning clinics, patients referred from general practicioners based on 1+ clinical risk factors for OP	Two QUS systems: CUBA Clinical (BUA, VOS), Sunlight Omnisense (distal radius, proximal phalanx mid finger, mid-shaft tibia)	DXA, LS-4, and total hip	Patients referred by general practitioner to DXA screening clinic	Unclear	Yes	Unclear	Unclear	Sample has potential for bias toward low BMD due to recruitment from DXA clinic (all patients referred by MD for clinical risk factors)

461
462

© 2018 American Medical Association. All rights reserved.

eTable 28. KQ 2a Risk of bias assessment forKQ 2a imaging studies predicting bone density status: Part 1, Participant Selection (continued)

First Author, Year	Patients (Setting, Intended Use of Index Test, Presentation, Prior Testing)	Index Test(s)	Reference Standard and Target Condition	Describe Methods of Patient Selection	Was A Consecutive or Random Sample of Patients Enrolled	Was a Case-Control Design Avoided?	Did the Study Avoid Inappropriate Exclusions?	Could the Selection of Patients Have Introduced Bias?	Comments
Harrison et al., 2006[26]	Caucasian females, 55-80 years (referred to clinical radiology, intended use of index test (QUS x2) underwent DXA and categorized as non-osteoporosis and osteoporosis. Subsequently underwent QUS and risk assessment using demographics and then combined algorithms-QUS used to predict osteoporosis	QUS x2	DXA	White Caucasian females aged 55 to 70 (mean 61, SD 4) years who were referred to Clinical Radiology, Imaging Science and Biomedical Engineering, University of Manchester for routine bone densitometry scans were invited to take part in the study	Unclear	Yes	Unclear	Low	No details on setting or how participants were selected
Jimenez-Nunez et al., 2013[27]	Women from primary and tertiary care, diagnosis, no prior testing	4 risk scores + PIXI of the heel	DXA of the hip and spine	Described as random from 2 sites	Yes	Yes	Yes	Low	NR

© 2018 American Medical Association. All rights reserved.

eTable 28. KQ 2a Risk of bias assessment forKQ 2a imaging studies predicting bone density status: Part 1, Participant Selection (continued)

First Author, Year	Patients (Setting, Intended Use of Index Test, Presentation, Prior Testing)	Index Test(s)	Reference Standard and Target Condition	Describe Methods of Patient Selection	Was A Consecutive or Random Sample of Patients Enrolled	Was a Case-Control Design Avoided?	Did the Study Avoid Inappropriate Exclusions?	Could the Selection of Patients Have Introduced Bias?	Comments
Kung et al., 2003[28]	Women in Hong Kong recruited from the community	OSTA index and QUI	DXA	Women from community, all comers who did not meet exclusion	Unclear	Yes	Yes	Low	Although noted to be early postmenopausal but ,eam age mean is 62
Kung et al., 2005[29]	Community of Asian (Southern Chinese) men; develop index based on clinical factors; compare clinical index with calcaneal QUS in predicting BMD (T< -2.5) by DXA	Clinical index	Calcaneal QUS; target condition - osteoporosis as determined by BMD at the hip and spine by DXA	Men from community, all comers who did not meet exclusion	Yes	Yes	Yes	Low	Unclear who chose to participate relative to larger group, excluded abnormal TSH group

© 2018 American Medical Association. All rights reserved.

eTable 28. KQ 2a Risk of bias assessment for KQ 2aimaging studies predicting bone density status: Part 1, Participant Selection (continued)

First Author, Year	Patients (Setting, Intended Use of Index Test, Presentation, Prior Testing)	Index Test(s)	Reference Standard and Target Condition	Describe Methods of Patient Selection	Was A Consecutive or Random Sample of Patients Enrolled	Was a Case-Control Design Avoided?	Did the Study Avoid Inappropriate Exclusions?	Could the Selection of Patients Have Introduced Bias?	Comments
Lynn et al., 2008[31]	US Caucasian (4658) and Hong Kong Chinese (1914) from the MrOS study with DXA and QUS measurements to compare screening tools (OST, MOST, QUI) to DXA	OST, MOST, QUI	DXA	US participants were recruited using population-based listings at six clinical settings in Birmingham, AL; Minneapolis, MN; Palo Alto, CA; Pittsburgh, PA; Portland, OR; and San Diego, CA. Hong Kong participants were recruited using a combination of private solicitation and public advertising from community centers, housing estates, and the general community. Men who had bilateral hip replacements or who were unable to walk without the assistance of another person were excluded.	Yes	Yes	Unclear	Low	NR

© 2018 American Medical Association. All rights reserved.

474 **eTable 28. KQ 2a Risk of bias assessment for KQ 2aimaging studies predicting bone density status: Part 1, Participant Selection**
475 **(continued)**

First Author, Year	Patients (Setting, Intended Use of Index Test, Presentation, Prior Testing)	Index Test(s)	Reference Standard and Target Condition	Describe Methods of Patient Selection	Was A Consecutive or Random Sample of Patients Enrolled	Was a Case-Control Design Avoided?	Did the Study Avoid Inappropriate Exclusions?	Could the Selection of Patients Have Introduced Bias?	Comments
McLeod et al., 2015[35]	Women referred for screening in Canada, no prior testing	QUS and OST	DXA	Patients referred for screening to one facility	Y es	Yes	Yes	Low	NA
Minnock et al., 2008[49]	Causian women underwent clinical risk factor questionnaire, QUS and DXA in order to determine whether a combined clinical assessment tool + QUS would be predictive of osteoporosis (low bone mass) by DXA	Combined clinical risk facors+ QUS	DXA	Women were referred to DXA scanning clinic at Great Western Hospital, Swindon, UK. Referral was performed by the patients GPs, or hospital based clinics	Unclear	Yes	Unclear	Unclear	Insufficient information

476
477

© 2018 American Medical Association. All rights reserved.

eTable 28. KQ 2a Risk of bias assessment for KQ 2aimaging studies predicting bone density status: Part 1, Participant Selection (continued)

First Author, Year	Patients (Setting, Intended Use of Index Test, Presentation, Prior Testing)	Index Test(s)	Reference Standard and Target Condition	Describe Methods of Patient Selection	Was A Consecutive or Random Sample of Patients Enrolled	Was a Case-Control Design Avoided?	Did the Study Avoid Inappropriate Exclusions?	Could the Selection of Patients Have Introduced Bias?	Comments
Richy et al., 2004[50]	Two cohorts of postmenopausal women, age 45 and older; purpose was to study #1 - develop an clinical algorithm tool+ QUS (n=407 women)with bone mass as the outcome measure, as derived from DXA, and then in study #2 used a second cohort (202 women) to validate the algorithm by comparing it to QUS alone and to the OST; community screening clinic; no prior testing	Clinical algorithm; QUS	DXA for low bone mass; osteoporosis	Women who attended public screening for Osteoporosis	Yes	Yes	Yes	Low	NR

© 2018 American Medical Association. All rights reserved.

eTable 28. KQ 2a Risk of bias assessment forKQ 2a imaging studies predicting bone density status: Part 1, Participant Selection (continued)

First Author, Year	Patients (Setting, Intended Use of Index Test, Presentation, Prior Testing)	Index Test(s)	Reference Standard and Target Condition	Describe Methods of Patient Selection	Was A Consecutive or Random Sample of Patients Enrolled	Was a Case-Control Design Avoided?	Did the Study Avoid Inappropriate Exclusions?	Could the Selection of Patients Have Introduced Bias?	Comments
Sinnott et al., 2006[45]	AA men, aged 35 and older (outpatient general medicine clinics at veteran hospital; intended use of clinical assessment tools and calcaneous ultrasound compared with the reference measure of BMD by DXA; no description of presentation in article; no prior testing): index text is ultrasound of calcaneous on non-dominant foot, outcome is low bone mass	ultrasound of calcaneous on non-dominant foot	BMD by DXA at the 1) lumbar spine (L1-L4) and 2) non-dominant hip(femoral neck, trochanter, total hip)	Participants were recruited from outpatient general medicine clinics at the Jesse Brown VA Medical Center over an 11-month period in 2004	Unclear	Yes	Yes	Low	Selection of participants may be a convenience sample but unclear. Men were recruited from general medicine clinics so selection bias likely low

Abbreviations: AL=Alabama; BMD= bone mineral density; BUA=broadband attenuation; CA=California; ; DXA=dual energy x-ray absorptiometry; GPs=general practitioners; KQ=key question; LS-4=lumbar spine 4; MD=medical doctor; MN=Minnesota; MOST=Male Osteoporosis Screening Tool; MrOS=Evaluation of osteoporosis screening tools for the osteoporotic fractures in men; NA=not applicable; OP=osteoporosis; OR=Oregon; OST=osteoporosis self-assessment tool; OSTA=Osteoporosis Self-assessment Tool for Asians; PA=Pennsylvania; QUI=ultrasound index; QUS=quantitative ultrasound; SD=standard deviation; TSH=thyroid-stimulating hormone; UK=United Kingdom; US=United States; VA=Veterans' Administration; VOS=velocity of sound.

© 2018 American Medical Association. All rights reserved.

eTable 29. KQ 2a Risk of Bias Assessment for KQ 2aimaging studies predicting bone density status: Part 2, Index Test

First Author, Year	Describe the Index Test and How It Was Conducted and Interpreted	Were the Index Test Results Interpreted Without Knowledge of the Results of the Reference Standard?	If A Threshold Was Used, Was It Pre-Specified?	Could the Conduct or Interpretation of the Index Test Have Introduced Bias?	Comments
Boonen et al., 2005[48]	QUS, DXR, RA	Yes	Yes	Low	NR
Cook et al., 2005[18]	Two QUS tests - CUBA clinical and Sunlight Omnisense measurements. Performed on non-dominant side with same ultrsaound gel. System quality verification tests each day.	Unclear	Yes	Unclear	Threshold question - yes and no used a 90% sensitivity threshold, but also created a cut off level based on the highest combined value of Sn and Sp. ROB assessment - depends on if QUS studies read independently of DXA imaging.
Harrison et al., 2006[26]	QUS x2	Unclear	Yes	Unclear	Osteoporosis status determined before index tests conducted, but unclear if results available
Jimenez-Nunez et al., 2013[27]	4 risk scores + PIXI of the heel, algorithms were developed	Yes	Yes	Low	NR
Kung et al., 2003[28]	Index characteristics through interview and QUI of right heel by technician	Unclear	Yes	Low	Index based on characteristics can be biased based on analysis decisions
Kung et al., 2005[29]	Index developed by authors based on characteristics	Unclear	Yes	Low	NR
Lynn et al., 2008[31]	OST, MOST, QUI	Unclear	Yes	Low	NR
McLeod et al., 2015[35]	QUS of BUA and SOS of left calcaneus & personal data based on questionnaire	Yes	Yes	Low	NR
Minnock et al., 2008[49]	Combined clinical risk facors+ QUS	Unclear	Yes	Low	NR
Richy et al., 2004[50]	Clinical algorithm; QUS	Unclear	Yes	Low	NR
Sinnott et al., 2006[45]	Ultrasound of calcaneous on non-dominant foot	Unclear	Unclear	Low	NR

Abbreviations: BUA=broadband attenuation; DXR=digital x-ray radiogrammetry; MOST=Male Osteoporosis Screening Tool; NR=not reported; OST=osteoporosis self-assessment tool; QUI=ultrasound index; QUS=quantitative ultrasound; RA=radiographic absorptiometry; Sn=sensitivity; SOS= speed of sound; Sp=specificity.

© 2018 American Medical Association. All rights reserved.

494 **eTable 30. KQ 2a Risk of bias assessment for KQ 2aimaging studies predicting bone density status: Part 3, Reference Test**

First Author, Year	Describe the Reference Standard and How It Was Conducted and Interpreted	Is the Reference Standard Likely to Correctly Classify the Target Condition?	Were the Reference Standard Results Interpreted Without Knowledge of the Results of the Index Test?	Could the Reference Standard, Its Conduct, or Its Interpretation Have Introduced Bias?	Comments
Boonen et al., 2005[48]	DXA, BMD of the lumbar spiine and proximal femur	Yes	Unclear	Low	NR
Cook et al., 2005[18]	DXA. BMD of the lumbar spine and total hip	Yes	Unclear	Unclear	NR
Harrison et al., 2006[26]	DXA, BMD of the femoral neck and total hip	Yes	Unclear	Low	NR
Jimenez-Nunez et al., 2013[27]	DXA, BMD of the hip and spine	Yes	Yes	Low	NR
Kung et al., 2003[28]	DXA: BMD of the lumbar spine, femoral neck	Yes	Unclear	Low	NR
Kung et al., 2005[29]	DXA: BMD of the lumbar spine, femoral neck	Yes	Yes	Low	NR
Lynn et al., 2008[31]	DXA, lumbar spine and proximal femur	Yes	Unclear	Low	All obtained from MrOS (sequence of data collection not described)
McLeod et al., 2015[35]	DXA: BMD of the lumbar spine, left and right femoral neck	Yes	Yes	Low	NR
Minnock et al., 2008[49]	DXA, BMD of the lumbar spine, femoral neck, and total hip	Yes	Unclear	Low	NR
Richy et al., 2004[50]	DXA, BMD of the femoral neck	Yes	Yes	Low	NR
Sinnott et al., 2006[45]	DXA; BMD of the hip, spine	Yes	Unclear	Low	NR

495 **Abbreviations:** BMD= bone mineral density; DXA=dual energy x-ray absorptiometry; KQ=key question; MrOS=Evaluation of osteoporosis screening tools for the osteoporotic
496 fractures in men; NR=not reported; QUS=quantitative ultrasound.

497

© 2018 American Medical Association. All rights reserved.

eTable 31. KQ 2a Risk of bias assessment for KQ 2aimaging studies predicting bone density status: Part 4, Participant Flow

First Author, Year	Describe Any Patients Who Did Not Receive the Index Test(S) and/or Reference Standard or Who Were Excluded	Describe the Time Interval and Any Interventions Between Index Test(S) and Reference Standard	Was There an Appropriate Interval Between Index Test(S) and Reference Standard?	Did All Patients Receive A Reference Standard?	Did Patients Receive The Same Reference Standard?	Were All Patients Included In The Analysis?
Boonen et al., 2005[48]	On treatment for osteo, peripheral oedema	Same day	Yes	Yes	Yes	Yes
Cook et al., 2005[18]	None	None	Yes	Yes	Yes	Yes
Harrison et al., 2006[26]	NR	NR	Unclear	Yes	Yes	Unclear
Jimenez-Nunez et al., 2013[27]	Nursing home, homebound, prior diagnosis of osteo, on osteo drugs, serious acute or chronic disease, hip replacement, steroids	Same day	Unclear	Yes	Yes	Unclear
Kung et al., 2003[28]	History or evidence of metabolic bone disease, menopause before 40, history of cancer, evidence of sig renal impariment, both hips previously fractured or replaced, prior use of any bisphosphonates, fluoride or calcitonin	NR	Unclear	Yes	Yes	Yes

© 2018 American Medical Association. All rights reserved.

eTable 31. KQ 2a Risk of bias assessment for KQ 2aimaging studies predicting bone density status: Part 4, Participant Flow (continued)

First Author, Year	Describe Any Patients Who Did Not Receive the Index Test(S) and/or Reference Standard or Who Were Excluded	Describe the Time Interval and Any Interventions Between Index Test(S) and Reference Standard	Was There an Appropriate Interval Between Index Test(S) and Reference Standard?	Did All Patients Receive A Reference Standard?	Did Patients Receive The Same Reference Standard?	Were All Patients Included In The Analysis?
Kung et al., 2005[29]	History or evidence of metabolic bone disease, hightory of cancer, evidence of sig renal impariment, both hips previously fractured or replaced, prior use of any bisphosphonates, fluoride or calcitonin, abnormal biochemisty including renal and liver function, serum calcium, phosphate, total alkaline phosphatase, and TSH	NR	Unclear	Yes	Yes	Yes
Lynn et al., 2008[31]	NR	NR	Unclear	Yes	Yes	NA
McLeod et al., 2015[35]	Previous diagnosis, progressive terminal illness	Within 3 weeks	Yes	Yes	Yes	Yes
Minnock et al., 2008[49]	NR	NR	Unclear	Yes	Yes	No
Richy et al., 2004[50]	NR	NR	Unclear	Yes	Yes	Yes
Sinnott et al., 2006[45]	NR	NR	Unclear	Yes	Yes	Yes

Abbreviations: NA=not applicable; NR=not reported; TSH=thyroid-stimulating hormone.

© 2018 American Medical Association. All rights reserved.

505 **eTable 32. KQ 2a Risk of bias assessment for imaging studies predicting bone density status: Part 5, Overall Ratings**

First Author, Year	Could the Patient Flow Have Introduced Bias?	Comments	Overall Judgement	Overall Comments
Boonen et al., 2005[48]	Low	NR	Low	Not a community-based sample. Women referred for bone densitometry.
Cook et al., 2005[18]	Low	NR	Unclear	Patient selection has the potential to skew the sample toward low BMD
Harrison et al., 2006[26]	Unclear	Participants underwent DXA and were categorized as non -osteo or osteo prior to QUS or risk indices	Unclear	Osteoporosis status determined first
Jimenez-Nunez et al., 2013[27]	Low	Random sample done with some sort of cards	Low	NR
Kung et al., 2003[28]	Low	NR	Low	NR
Kung et al., 2005[29]	Low	It is not clear what the time frame between clinical assessment of risk factors and QUS; however should be little impact;I put that all participants received the same reference standard (referring to the validated group here)	Low	NR
Lynn et al., 2008[31]	Low	NR	Low	NR
McLeod et al., 2015[35]	Low	Effort made to contact patient, enroll and conduct OST and QUS within 3 weeks of DXA scan to complete study assessments prior to provider receiving DXA results and talking with patient.	Low	NR
Minnock et al., 2008[49]	Low	NR	Unclear	Initial sample is 274 but number in analysis is 235 because of missing data, impact of missing data unclear
Richy et al., 2004[50]	Low	NR	Low	NR
Sinnott et al., 2006[45]	Low	The flow was not specifically described, but appears sequence was clinical assessment followed by ultrasound and then DXA.	Low	NR

506 **Abbreviations:** BMD=body mineral density; DXA=dual energy x-ray absorptiometry; KQ=key question; MrOS=Evaluation of osteoporosis screening tools for the osteoporotic
507 fractures in men; NR=not reported; OST=osteoporosis self-assessment tool; QUS=quantitative ultrasound.

508

© 2018 American Medical Association. All rights reserved.

509 **eTable 33. KQ 2a Fracture prediction studies risk of bias: Part 1, Participant Selection**

First Author, Year	Describe Screening or Treatment Interventions and Comparators	Prediction Model Development as Well as Testing of Predictive Performance in Other Individuals (External Validation), Both in the Same Publication	Testing the Predictive Performance of a Previously Developed Prediction Model in Other Individuals (External Validation)	For Validity Were Appropriate Data Sources Used, e.g. Cohort, RCT or Nested Case-Control Study Data?
Ahmed, 2014[51]	1. Garvan FRC with BMD, adjusted for age, prior fracture, prior fall 2. Garvan FRC, adjusted for body weight, age, prior fracture, prior fall.	No	Yes- Val only	Yes
Azagra, 2011[52]	FRAX (Spain)	No	No	Probably no
Bauer, 2007[53]	Quantitative US	No	No	Yes
Berry, 2013[54]	Assess contribution of repeat BMD in 4 years to fx risk : 1. BMD at basline and Fx risk 2. BMD percent change and Fx risk 3. BMD at baseline and BMD Percent change and Fx risk	No	Yes- Val only	Yes
Chan, 2012[55]	1. FNBMD (adjusted for age, falls, prior fracture) 2. QUS (BUA) plus FNBMD (adjusted for age, falls, prior fracture)	No	Yes- Val only	Yes
Chan, 2013[56]	1. FNplus BMD (adjusted for age, falls, prior fracture) 2. QUS (BUA) plus FNBMD (adjusted for age, falls, prior fracture)	No	Yes- Val only	Yes
Crandall, 2014[57]	Comparison of three screening strategies for women age 50-64: 1. USPSTF Strategy (FRAX 3.0 without BMD, with followup BMD testing for fx risk >= 9.3%)- 10 yr horizon 2. OST-horizon unknown, this was developed to identify osteoporosis, not fracture 3. SCORE-horizon unknown, this was developed to identify osteoporosis, not fracture	No	Yes- Val only	Yes

© 2018 American Medical Association. All rights reserved.

eTable 33. KQ 2a Fracture prediction studies risk of bias: Part 1, Participant Selection (continued)

First Author, Year	Describe Screening or Treatment Interventions and Comparators	Prediction Model Development as Well as Testing of Predictive Performance in Other Individuals (External Validation), Both in the Same Publication	Testing the Predictive Performance of a Previously Developed Prediction Model in Other Individuals (External Validation)	For Validity Were Appropriate Data Sources Used, e.g. Cohort, RCT or Nested Case-Control Study Data?
Hans, 2011[58]	TBS alone, DXA alone, TBS plus DXA	No	No	Probably yes
Hillier, 2007[59]	Imaging screening: DXA, initial BMD, repeat BMD, change in BMD, initial BMD plus change in BMD	No	No	Yes
Hippisley-Cox, 2012[60]	Qfracture updated with additional clinical predictors and outcomes	Yes- Dev and Val	Yes- Val only	Yes
Iki, 2014[61]	DXA - spine areal BMD, trabecular bone score	No	No	Yes
Iki, 2015 [62]	FRAX and TBS	no	Yes- Val only	yes
Kalveston, 2016 [63]	FRAX and BMD	Yes- Val only	Yes- Val only	Yes
Kanis, 2007[64]	FRAX	Yes- Dev and Val	No	Yes
Kwok, 2012[65]	Imaging screening: QUS (BUA, SOS, QUI measures), DXA (tHIP, fnHIP, spine BMD)	No	No	Yes
Leslie, 2010[66]	CAROC	No	Yes	Yes
Leslie, 2012[67]	FRAX	No	Yes- Dev and Val	Yes
Leslie, 2012[68]	FRAX with and without DXA	No	Yes	Yes
Leslie, 2013[69]	Trabecular bone score	No	No	Yes
Lo, 2011[70]	FRC	No	Yes- Val only	Probably yes
Lundin, 2015 [71]	FRAX and BMD	no	Yes- Val only	yes
Melton, 2005[72]	NOF model including femoral neck BMD and clinical risk factors (personal fx history, FHx, low BWT, smoking status)	No	Yes- Val only	Yes
Miller, 2002[73]	Heel SXR, Heel QUS, forearm DXA, finger DXA; NORA study	No	No	Yes
Morin, 2009[36]	body weight, BMI, OST	No	Yes	Yes
Nguyen, 2004[74]	QUS. DOES	No	No	Yes
Rubin, 2013[75]	FRAX (no BMD), OST, ORAI, OSIRIS, SCORE, Age alone	No	Yes- Val only	Yes

© 2018 American Medical Association. All rights reserved.

eTable 33. KQ 2a Fracture prediction studies risk of bias: Part 1, Participant Selection (continued)

First Author, Year	Describe Screening or Treatment Interventions and Comparators	Prediction Model Development as Well as Testing of Predictive Performance in Other Individuals (External Validation), Both in the Same Publication	Testing the Predictive Performance of a Previously Developed Prediction Model in Other Individuals (External Validation)	For Validity Were Appropriate Data Sources Used, e.g. Cohort, RCT or Nested Case-Control Study Data?
Stewart, 2006[76]	DXA	No	Yes- Val only	Yes
van Geel, 2014[77]	FRAX, Garvan FRCr	No	Yes- Val only	Yes

Abbreviations: BMD= bone mineral density; BMI=body mass index; BUA=broadband attenuation; BWT=body weight; CAROC=Canadian Association of Radiologists and Osteoporosis Canada; DOES=Dubbo Osteoporosis Epidemiology Study; DXA=dual energy x-ray absorptiometry; FNBMD=femoral neck bone mineral density; fnHIP=femoral neck of hip; FNplus=femoral neck plus; FRAX=Fracture Risk Assessment tool; FRC=Fracture Risk Calculator; Fx=fracture; NOF=National Osteoporosis Foundation; NORA=National Osteoporosis Risk Assessment; ORAI=Osteoporosis Risk Assessment Instrument; OSIRIS=Osteoporosis Index of Risk; OST=osteoporosis self-assessment tool; QUI=ultrasound index; QUS=quantitative ultrasound; SCORE=Simple Calculated Osteoporosis Risk Estimation Tool; SOS=speed of sound; SXR=single x-ray absorptiometry; TBS=trabecular bone score; tHIP=total hip; US=United States; USPSTF=United States Preventive Services Task Force.

© 2018 American Medical Association. All rights reserved.

521 **eTable 34. KQ 2a Fracture prediction studies risk of bias: Part 2, Participant Selection**

First Author, Year	For Validity Were all Inclusions And Exclusions of Participants Appropriate?	For Validity Were Participants Enrolled at a Similar State of Health, or Were Predictors Considered to Account for Any Dissimilarities?	Risk of Bias Introduced by Selection of Participants	Justification of Bias Rating	Comments
Ahmed, 2014[51]	Yes	Yes	Low	NR	NR
Azagra, 2011[52]	Yes	Yes	Unclear	Cohort was assembled from participants referred for screening by primary or specialty care physicians. Thus, the cohort does not represent an entirely unselected population.	NR
Bauer, 2007[53]	Yes	Yes	Low	NR	NR
Berry, 2013[54]	Yes	Yes	Low	NR	NR
Chan, 2012[55]	Yes	Yes	Low	NR	NR
Chan, 2013[56]	No	Yes	High	High concern for spectrum bias in the subgroup analysis, since participants in the analysis were limited to those with BMD < -2.5	NR
Crandall, 2014[57]	Yes	Yes	Low	NR	NR
Hans, 2011[58]	Probably yes	Probably yes	Low	NR	NR
Hillier, 2007[59]	Probably yes	Yes	Low	NR	NR
Hippisley-Cox, 2012[60]	Probably yes	Probably yes	Low	NR	NR
Iki, 2014[61]	Yes	Yes	Low	NR	NR
Iki, 2015[62]	yes	yes	low	Population-based cohort	None
Kalvesten, 2016[63]	yes	yes	low	Population-based recruitment into study.	None
Kanis, 2007[64]	No information	Probably yes	Low	NR	Inclusion/exclusion criteria for the 11 independent validation cohorts is not included.
Kwok, 2012[65]	Yes	Yes	Low	NR	NR

522
523

© 2018 American Medical Association. All rights reserved.

524 **eTable 34. KQ 2a Fracture prediction studies risk of bias: Part 2, Participant Selection (continued)**

First Author, Year	For Validity Were all Inclusions And Exclusions of Participants Appropriate?	For Validity Were Participants Enrolled at a Similar State of Health, or Were Predictors Considered to Account for Any Dissimilarities?	Risk of Bias Introduced by Selection of Participants	Justification of Bias Rating	Comments
Leslie, 2010[66]	No information	Probably no	Low	Database covers population in Manitoba aged 50 with a first bone desnity measurement, and all citizens of Manitoba have university access to publicly funded medical care including BMD.	NR
Leslie, 2012[67]	No information	Probably no	Low	Database covers population in Manitoba aged 50 with a first bone desnity measurement, and all citizens of Manitoba have university access to publicly funded medical care including BMD.	NR
Leslie, 2012[68]	No information	Probably no	Low	Database covers population in Manitoba aged 50 with a first bone desnity measurement, and all citizens of Manitoba have university access to publicly funded medical care including BMD.	NR
Leslie, 2013[69]	No information	Probably no	Low	Database covers all women in Manitoba aged 50 with a first bone desnity measurement, and all citizens of Manitoba have university access to publicly funded medical care including BMD.	NR

525
526

© 2018 American Medical Association. All rights reserved.

eTable 34. KQ 2a Fracture prediction studies risk of bias: Part 2, Participant Selection (continued)

First Author, Year	For Validity Were all Inclusions And Exclusions of Participants Appropriate?	For Validity Were Participants Enrolled at a Similar State of Health, or Were Predictors Considered to Account for Any Dissimilarities?	Risk of Bias Introduced by Selection of Participants	Justification of Bias Rating	Comments
Lo, 2011[70]	Probably no	Probably yes	Unclear	Possible spectrum bias due to use of population of women referred for DXA testing. Other exclusions may also have introduced some selection bias. Impact of these cannot be determined. Only about 94,000 of an eligible population of 500,000 were analyzed.	Study limited to women aged 50 to 85 who were referred to have bone density scanning. Women without continuous membership both prior and following DXA scans, and those for whom DXA results were not electronically accessible and those with missing race/eth
Lundin, 2015[71]	Yes	Yes	Low	Population based recruitment strategy.	None
Melton, 2005[72]	No information	No information	Unclear	NR	NR
Miller, 2002[73]	Yes	No information	Unclear	It is unclear whether sites selected people with similar underlying characteristics.	NR
Morin, 2009[36]	No information	Probably no	Low	Database covers all women in Manitoba aged 40 to 59 with a first bone desnity measurement, and all citizens of Manitoba have university access to publicly funded medical care including BMD.	NR
Nguyen, 2004[74]	No information	No information	Unclear	unclear whether patients selected from database similar underlying characteristics	NR
Rubin, 2013[75]	Yes	Yes	Low	NR	NR
Stewart, 2006[76]	Yes	No information	Low	NR	NR
van Geel, 2014[77]	Probably yes	Yes	Low	NR	NR

Abbreviations: BMD= bone mineral density; DXA=dual energy x-ray absorptiometry; KQ=key question; NR=not reported.

© 2018 American Medical Association. All rights reserved.

eTable 35. KQ 2a Fracture prediction studies risk of bias: Part 3, Predictors

First Author, Year	For Validity Were Predictors Defined and Assessed in a Similar Way for All Participants?	For Validity Were Predictors Defined and Assessed in a Similar Way to Predictors in the Development Model?	Risk of Bias Introduced by Predictors or Their Assessment	Justification of Bias Rating	Comments
Ahmed, 2014[51]	Yes	yes	Low	NR	NR
Azagra, 2011[52]	Yes	Yes	Low	NR	NR
Bauer, 2007[53]	Yes	Yes	YesLow	NR	NR
Berry, 2013[54]	Yes	Yes	Low	NR	NR
Chan, 2012[55]	Yes	Yes	Low	NR	NR
Chan, 2013[56]	Yes	Yes	Low	NR	NR
Crandall, 2014[57]	Yes	Yes for FRAX and OST, probably no for SCORE	Low	Authors show that use of different age cut off for prior history of fracture would likely have little impact.	NR
Hans, 2011[58]	NA-NOT VAL	NA-NOT VAL	Low	NR	NR
Hillier, 2007[59]	NA-NOT VAL	NA-NOT VAL	Low	NR	NR
Hippisley-Cox, 2012[60]	Yes	Yes	Low	NR	NR
Iki, 2014[61]	Yes	NA-NOT VAL	Low	NR	NR
Iki, 2015 [62]	yes	yes	low	In person interviews	None
Kalvesten, 2016 [63]	yes	yes	low	Questionnaire-based assessment, all relevant predictors assessed.	None
Kanis, 2007[64]	Probably yes	Probably yes	Low	NR	NR
Kwok, 2012[65]	NA-NOT VAL	NA-NOT VAL	Low	Imaging prediction of fracture - not clinical prediction tool	NR
Leslie, 2010[66]	Yes	No	Unclear	The final risk category was modified to reflect the presence of additional risk factors: any prior osteoporotic fracture (from 1987 to the date of BMD testing) and/or recent systemic corticosteroid use (in the year before BMD testing).	NR

© 2018 American Medical Association. All rights reserved.

eTable 35. KQ 2a Fracture prediction studies risk of bias: Part 3, Predictors (continued)

First Author, Year	For Validity Were Predictors Defined and Assessed in a Similar Way for All Participants?	For Validity Were Predictors Defined and Assessed in a Similar Way to Predictors in the Development Model?	Risk of Bias Introduced by Predictors or Their Assessment	Justification of Bias Rating	Comments
Leslie, 2012[67]	Yes	No	Unclear	Parental hip fracture information missing for FRAX probability estimates prior to 2005, adjusted using age- and sex-specific adjustment factors derived from 2005 to 2008 parental hip fracture responses	NR
Leslie, 2012[68]	Yes	No	Unclear	Parental hip fracture information missing for FRAX probability estimates prior to 2005, adjusted using age- and sex-specific adjustment factors derived from 2005 to 2008 parental hip fracture responses	NR
Leslie, 2013[69]	Yes	NA	Low	TBS assessed the same way for all	NR
Lo, 2011[70]	Yes	Probably yes	Low	NR	NR
Lundin, 2015[71]	Yes, for DXA No, for FRAX	Yes, for DXA No information, for FRAX	low for DXA unclear for FRAX	The study does not describe how inputs to FRAX were obtained	NR
Melton, 2005[72]	Yes	Probably yes	Low	NR	NR
Miller, 2002[73]	Yes	NA	Low	Peripheral bone densitometry done in simiarl ways for all	NR
Morin, 2009[36]	Yes	No information	Unclear	Unclear whether data for OST (age, weight) was collected before fracture for all participants	NR
Nguyen, 2004[74]	Yes	NA	NALow	QUS done in similar ways for all	NR

© 2018 American Medical Association. All rights reserved.

eTable 35. KQ 2a Fracture prediction studies risk of bias: Part 3, Predictors (continued)

First Author, Year	For Validity Were Predictors Defined and Assessed in a Similar Way for All Participants?	For Validity Were Predictors Defined and Assessed in a Similar Way to Predictors in the Development Model?	Risk of Bias Introduced by Predictors or Their Assessment	Justification of Bias Rating	Comments
Rubin, 2013[75]	Yes	No information	Low	NR	NR
Stewart, 2006[76]	Yes	Yes	Low	NR	NR
van Geel, 2014[77]	Probably yes	Probably yes	Low	NR	NR

Abbreviations: BMD= bone mineral density; BMI=body mass index; BUA=broadband attenuation; BWT=body weight; CAROC=Canadian Association of Radiologists and Osteoporosis Canada; DOES=Dubbo Osteoporosis Epidemiology Study; DXA=dual energy x-ray absorptiometry; FHx=fracture history; FNBMD=femoral neck BMD; fnHIP=femoral neck of hip; FNplus=femoral neck plus; FRAX=Fracture Risk Assessment tool; FRC=Fracture Risk Calculator; Fx=fracture; KQ=key question; NOF=National Osteoporosis Foundation; NORA=National Osteoporosis Risk Assessment; NR=not reported; ORAI=Osteoporosis Risk Assessment Instrument; OSIRIS=Osteoporosis Index of Risk; OST=osteoporosis self-assessment tool; QUI=ultrasound index; QUS=quantitative ultrasound; SCORE=Simple Calculated Osteoporosis Risk Estimation Tool; SOS=speed of sound; SXR=single x-ray absorptiometry; TBS=trabecular bone score; tHIP=total hip; US=United States; USPSTF=United States Preventive Services Task Force; VAL=validity.

© 2018 American Medical Association. All rights reserved.

eTable 36. KQ 2a Fracture prediction studies risk of bias: Part 4, Outcomes

First Author, Year	For Validity Was a Pre-Specified Outcome Definition Used?	For Validity Was the Outcome Defined and Determined in a Similar Way for All Participants?	For Validity Was the Outcome Defined and Determined in a Similar Way to the Outcome in the Development Model?	For Validity Was the Outcome Determined Without Knowledge of predictor Information?
Ahmed, 2014[51]	Yes	Yes	Yes	No information
Azagra, 2011[52]	Yes	Yes	Yes	Yes
Bauer, 2007[53]	Yes	Yes	Yes	No information
Berry, 2013[54]	Yes	Yes	Yes	No information
Chan, 2012[55]	Yes	Yes	Probably yes	No information
Chan, 2013[56]	Yes	Yes	Probably yes	No information
Crandall, 2014[57]	Yes	Yes	No for OST and SCORE, Yes for FRAX	No information
Hans, 2011[58]	Yes	Yes	NA-NOT VAL	Yes
Hillier, 2007[59]	Yes	Yes	NA-NOT VAL	Yes
Hippisley-Cox, 2012[60]	Yes	Yes	Yes	Yes
Iki, 2014[61]	Yes	Yes	NA-NOT VAL	Yes
Iki, 2015[62]	yes	yes	yes	no information
Kalvesten, 2016[63]	yes	yes	yes	no information
Kanis, 2007[64]	No information	No	Probably yes	No information
Kwok, 2012[65]	Yes	Yes	NA-NOT VAL	Yes
Leslie, 2010[66]	Yes	Yes	No information	Probably yes
Leslie, 2012[67]	Yes	Yes	No information	Probably yes
Leslie, 2012[68]	Yes	Yes	No information	Probably yes
Leslie, 2013[69]	Yes	Yes	Yes	Yes
Lo, 2011[70]	Yes	Yes	Probably yes	No information
Lundin, 2015[71]	yes	yes	yes	no information
Melton, 2005[72]	Yes	Yes	Probably no	Yes
Miller, 2002[73]	Yes	Yes	Yes	Yes
Morin, 2009[36]	Yes	Yes	No information	No information
Nguyen, 2004[74]	Yes	Yes	Yes	Yes
Rubin, 2013[75]	Yes	Yes	No information	Yes
Stewart, 2006[76]	Yes	Yes	Yes	No information
van Geel, 2014[77]	Yes	Yes	Probably yes	Yes

Abbreviations: FRAX=Fracture Risk Assessment tool; KQ=key question; OST=osteoporosis self-assessment tool; SCORE=Simple Calculated Osteoporosis Risk Estimation Tool; VAL=validity.

© 2018 American Medical Association. All rights reserved.

548 **eTable 37. KQ 2a Fracture prediction studies risk of bias: Part 5, Outcomes**

First Author, Year	Risk of Bias Introduced by the Outcome or Its Determination	Justification of Bias Rating	Comments
Ahmed, 2014[51]	Low	NR	NR
Azagra, 2011[52]	Low	NR	NR
Bauer, 2007[53]	Low	NR	NR
Berry, 2013[54]	Low	NR	NR
Chan, 2012[55]	Low	NR	NR
Chan, 2013[56]	Low	NR	NR
Crandall, 2014[57]	Unclear	NR	Both OST and SCORE were initially developed and validated for prediction of low BMD. In this study they are being used to predict fracture. It's unclear what impact this will have.
Hans, 2011[58]	Low	NR	NR
Hillier, 2007[59]	Low	NR	NR
Hippisley-Cox, 2012[60]	Low	NR	NR
Iki, 2014[61]	Low	NR	NR
Iki, 2015[62]	Low	fractures were confirmed	None
Kalvesten, 2016[63]	Low	Confirmation of all self-reported fractures. Outcomes censored at 10 years.	NR
Kanis, 2007[64]	Unclear	Fracture ascertainment was by self-report in some cohorts and by medical record or radiology report confirmation in other cohorts.	NR
Kwok, 2012[65]	Low	Did not exclude traumatic fractures; would just have to take fragility fracture #s	NR
Leslie, 2010[66]	Low	NR	NR
Leslie, 2012[67]	Low	NR	NR
Leslie, 2012[68]	Low	NR	NR
Leslie, 2013[69]	Low	NR	NR
Lo, 2011[70]	Low	NR	NR
Lundin, 2015[71]	Low	Identification of fractures from population based claims/diagnosis data.	None
Melton, 2005[72]	High	13.3% fractures were due to severe trauma, another 18.3% unclear cause	NR
Miller, 2002[73]	High	self-reported factures	NR
Morin, 2009[36]	Unclear	unclear whether OST variables collected for all women before fracture outcome	NR
Nguyen, 2004[74]	Low	NR	NR

549

© 2018 American Medical Association. All rights reserved.

eTable 37. KQ 2a Fracture prediction studies risk of bias: Part 5, Outcomes (continued)

First Author, Year	Risk of Bias Introduced by the Outcome or Its Determination	Justification of Bias Rating	Comments
Rubin, 2013[75]	Low	NR	NR
Stewart, 2006[76]	Low	NR	NR
van Geel, 2014[77]	Low	NR	NR

Abbreviations: BMD= bone mineral density; KQ=key question; NR=not reported; OST=osteoporosis self-assessment tool; SCORE=Simple Calculated Osteoporosis Risk Estimation Tool.

© 2018 American Medical Association. All rights reserved.

555 **eTable 38. KQ 2a Fracture prediction studies risk of bias: Part 6, Participant Flow**

First Author, Year	Describe Missing Data on Predictors and Outcomes as Well as Methods Used for Missing Data (What Was The Missing Data, how was it managed? Focus on Diff Between N Eligible and N Analyzed, Other Data Issues)	For Validity Were There a Reasonable Number of Outcome Events? (Bring up for Discussion if Low)	For Validity Was the time interval between Predictor Assessment and Outcome Determination Appropriate?	For Validity Were all Enrolled Participants Included in the analysis? If over 50% or Otherwise Concerning, Bring up for Discussion
Ahmed, 2014[51]	Participants with missing data were excluded.	Yes	Yes for 5 years, No for 10 years	Yes
Azagra, 2011[52]	Not clear how missing data handled.	Yes	Yes	No
Bauer, 2007[53]	No missing data.	Yes	Yes	Yes
Berry, 2013[54]	No data on parental history of hip fracture, set to "no".	Yes	Yes	Yes
Chan, 2012[55]	No missing data described	Yes	Yes	Probably no
Chan, 2013[56]	No missing data described	Yes	Yes	Probably no
Crandall, 2014[57]	Missing data set to "not present". Most common predictor missing was parental hip fx history.	Yes	Yes	Yes
Hans, 2011[58]	N eligble NR (over 34,000, see comments) N included 29407	Yes	Probably yes	Probably yes
Hillier, 2007[59]	9704 enrolled in SOF, 8141 women had followup (93%), 4124 had repeat BMD measurement, excluded patients with incident hip or non-spine fractures between BMD measurement (72, 513 respectively)	Yes	Yes	Probably no
Hippisley-Cox, 2012[60]	Did not report amount of missing data (particularly for BMI, smoking Status, alcohol intake), though report mulitple imputation was used. Qresearch database > 13,000,000 patients but only 4,726,046 used for development and validation cohorts. Only inclusi	Yes	Probably yes	No

556

© 2018 American Medical Association. All rights reserved.

eTable 38. KQ 2a Fracture prediction studies risk of bias: Part 6, Participant Flow (continued)

First Author, Year	Describe Missing Data on Predictors and Outcomes as Well as Methods Used for Missing Data (What Was The Missing Data, how was it managed? Focus on Diff Between N Eligible and N Analyzed, Other Data Issues)	For Validity Were There a Reasonable Number of Outcome Events? (Bring up for Discussion if Low)	For Validity Was the time interval between Predictor Assessment and Outcome Determination Appropriate?	For Validity Were all Enrolled Participants Included in the analysis? If over 50% or Otherwise Concerning, Bring up for Discussion
Iki, 2014[61]	789 eligible 665 analyzed 112 lost to followup 4 unassessable VFA 8 developed disease affecting bone metabolism	Yes	Yes	Probably yes
Iki, 2015[62]	No information about the men excluded from the analysis.	probably no	probably no	probably yes
Kalvesten, 2016[63]	Only participants with complete data were included in analysis.	yes	yes	probably no
Kanis, 2007[64]	Sensitivity analyses used to assess impact of missing predictor information.	Probably yes	Yes	No information
Kwok, 2012[65]	N (eligible)=2000, N (analyzed)=1921, those missing QUS or DXA readings excluded, invalid QUS readings excluded	Probably yes	Probably yes	No
Leslie, 2010[66]	Unclear	Yes	Yes	Yes
Leslie, 2012[67]	Unclear	Yes	Yes	Yes
Leslie, 2012[68]	Unclear	Yes	Yes	Yes
Leslie, 2013[69]	NR	Yes	Yes	Probably yes
Lo, 2011[70]	Women with missing data on race/ethnicity and BMD were excluded from analysis.	Yes	Yes	Yes
Lundin, 2015[71]	Missing data for 5 participants	yes	yes	yes
Melton, 2005[72]	1,479 approached, 1,315 eligible, 655 consented, only 393 included in analysis - unclear why	Probably yes	Yes	No
Miller, 2002[73]	NR	Yes	No	Unclear
Morin, 2009[36]	NR	Yes	Yes	Unclear
Nguyen, 2004[74]	NR	Yes	Unclear	Unclear

© 2018 American Medical Association. All rights reserved.

eTable 38. KQ 2a Fracture prediction studies risk of bias: Part 6, Participant Flow (continued)

First Author, Year	Describe Missing Data on Predictors and Outcomes as Well as Methods Used for Missing Data (What Was The Missing Data, how was it managed? Focus on Diff Between N Eligible and N Analyzed, Other Data Issues)	For Validity Were There a Reasonable Number of Outcome Events? (Bring up for Discussion if Low)	For Validity Was the time interval between Predictor Assessment and Outcome Determination Appropriate?	For Validity Were all Enrolled Participants Included in the analysis? If over 50% or Otherwise Concerning, Bring up for Discussion
Rubin, 2013[75]	Eligible: 5000 Analysis: 3614 Exclusion: 334 missing questionnaire data, 246 diagnosed with/treated for OP, reported "near complete follow-up" in registry	Probably yes	Probably no	Yes
Stewart, 2006[76]	Nonresponse analysis done.	Yes	Yes	Yes
van Geel, 2014[77]	Random sample: 1686, analysis sample: 506 Missing: no coop with MD (272), no coop with patient (448), untraceable/deceased (207), age <60 (110)	Probably yes	Probably no	Yes

Abbreviations: BMD= bone mineral density; BMI=body mass index; DXA=dual energy x-ray absorptiometry; KQ=key question; MD=medical doctor; N=number; NR=not reported; OP=osteoporosis; QUS=quantitative ultrasound; SOF=study of osteoporotic fractures; VFA=vertebral fracture assessment.

© 2018 American Medical Association. All rights reserved.

564 **eTable 39. KQ 2a Fracture prediction studies risk of bias: Part 7, Participant Flow**

First Author, Year	For Validity Were Participants With Missing Data Handled Appropriately?	Risk of Bias Introduced By Sample Size or Participant Flow	Justification of Bias Rating	Comments
Ahmed, 2014[51]	Yes	Low for 5 yr outcomes; unclear for 10 yr outcomes	Inadequate duration of follow-up for 10 year risk predictions.	NR
Azagra, 2011[52]	No information	Unclear	Unclear eligible N	NR
Bauer, 2007[53]	Yes	Low	NR	No mention of missing data
Berry, 2013[54]	Yes	Low	NR	NR
Chan, 2012[55]	Yes	Unclear	Some members of the cohort began before the use of QUS was introduced, so they would not be eligible. It's still not clear why of the 3678 eligible in the cohort, < 1,000 comprised the analytic sample	NR
Chan, 2013[56]	Yes	Unclear	NR	NR
Crandall, 2014[57]	Yes	Low	NR	NR
Hans, 2011[58]	Probably yes	Low	NR	No mention of missing data Only says matching of personal identifier information with the administrative data repository in over 34,000 DXA patients was achieved in over 99%
Hillier, 2007[59]	Yes	Low	NR	NR
Hippisley-Cox, 2012[60]	Probably yes	Unclear	Unclear exclusion criteria	Over 13 million in database, only 4.7 million used
Iki, 2014[61]	Probably yes	Low	NR	NR
Iki, 2015[62]	Probably yes	Unclear	Follow-up was only 4.5 yrs, but using a 10 year risk prediction. 93% of those enrolled were included in the analysis.	NR

565
566

© 2018 American Medical Association. All rights reserved.

567 **eTable 39. KQ 2a Fracture prediction studies risk of bias: Part 7, Participant Flow (continued)**

First Author, Year	For Validity Were Participants With Missing Data Handled Appropriately?	Risk of Bias Introduced By Sample Size or Participant Flow	Justification of Bias Rating	Comments
Kalvesten, 2016[63]	Probably yes	Unclear	The entire study cohort was about 9000, but not all had complete data for calculation of FRAX and DXA measurement. Thus, analysis restricted to those with complete data, those included were younger and a little healthier and had lower prevalence of prior fracture; though BMD measures were similar.	NR
Kanis, 2007[64]	Probably yes	Low	NR	NR
Kwok, 2012[65]	Probably yes	Low	2.5% excluded for missing data (small)	NR
Leslie, 2010[66]	No information	Unclear	NR	NR
Leslie, 2012[67]	No information	Unclear	NR	NR
Leslie, 2012[68]	No information	Unclear	NR	NR
Leslie, 2013[69]	Probably yes	Low	99% accuracy and completeness	NR
Lo, 2011[70]	Probably yes	Low	NR	NR
Lundin, 2015[71]	Yes	Low	No concerns	NR
Melton, 2005[72]	No information	High	Only about 50% of eligible patietns consented, and of those only 2/3rd included for analysis	NR
Miller, 2002[73]	No information	Unclear	Unclear whether followup window is sufficient	NR
Morin, 2009[36]	No information	Unclear	Unclear what proportion of cohort did not have information on predictors	NR
Nguyen, 2004[74]	No information	Unclear	The average time between imaging and fractures is unclear	NR
Rubin, 2013[75]	No information	Unclear	Only 3 year follow-up while FRAX predicts 10 year fracture for women over 40 years old	NR

568
569

© 2018 American Medical Association. All rights reserved.

570 **eTable 39. KQ 2a Fracture prediction studies risk of bias: Part 7, Participant Flow (continued)**

First Author, Year	For Validity Were Participants With Missing Data Handled Appropriately?	Risk of Bias Introduced By Sample Size or Participant Flow	Justification of Bias Rating	Comments
Stewart, 2006[76]	Probably yes	Low	NR	NR
van Geel, 2014[77]	Probably yes	Unclear	FRAX and Garvan predict 10 year risk - follow-up only for 5 years. Likely underestimates risk. 124 of 630 patients lost to follow-up	NR

571 **Abbreviations:** DXA=dual energy x-ray absorptiometry; FRAX=Fracture Risk Assessment tool; KQ=key question; N=number; NR=not reported; QUS=quantitative ultrasound.
572
573

© 2018 American Medical Association. All rights reserved.

eTable 40. KQ 2a Fracture prediction studies risk of bias: Part 8, Analysis

First Author, Year	For Validity Were Non-Binary Predictors Handled Appropriately?	For Validity Were any Complexities in the Data (e.g. Competing Risks, Multiple Events Per Individual) Accounted for Appropriately?	For Validity Was the Model Recalibrated or was it Likely (based on the Evidence Presented, e.g. Calibration Plot) That Recalibration Was Not Needed?	Risk of Bias Introduced by the Analysis
Ahmed, 2014[51]	Probably yes	No information	Probably no	Unclear for AUC High for NRIs at both 5 and 10 yrs.
Azagra, 2011[52]	Yes	Probably yes	Yes	Low
Bauer, 2007[53]	Yes	No information	No information	Low
Berry, 2013[54]	Yes	No information	Probably yes	Low
Chan, 2012[55]	Probably yes	No information	Yes	Varies by outcome
Chan, 2013[56]	Probably yes	No information	Yes	Varies by outcome
Crandall, 2014[57]	Yes	No information	No information	Unclear
Hans, 2011[58]	Yes	Probably yes	Probably yes	Low
Hillier, 2007[59]	Probably yes	Yes	Yes	Low
Hippisley-Cox, 2012[60]	Probably yes	Yes	Yes	Low
Iki, 2014[61]	Yes	Probably yes	Yes	Low
Iki, 2015[62]	Yes	No information	Yes	Low
Kalvesten, 2016[63]	Yes	No information	Yes	Low
Kanis, 2007[64]	Yes	Probably yes	Probably yes	Low
Kwok, 2012[65]	Yes	Yes	NA-NOT VAL	Low
Leslie, 2010[66]	Yes	No information	No	Low
Leslie, 2012[67]	Yes	No information	No	Low
Leslie, 2012[68]	Yes	No information	No	Low
Leslie, 2013[69]	Yes	No information	No	Low
Lo, 2011[70]	Yes	No information	Probably yes	Low
Lundin, 2015[71]	Yes	No	Yes	Low
Melton, 2005[72]	Yes	Probably yes	Yes	Low
Miller, 2002[73]	Yes	No information	No	Low
Morin, 2009[36]	Yes	No information	No	Low
Nguyen, 2004[74]	Yes	No information	No	Low
Rubin, 2013[75]	Yes	Yes	Yes	Low
Stewart, 2006[76]	NA	Probably no	Na	Low
van Geel, 2014[77]	Yes	Probably yes	Yes	Low

Abbreviations: AUC= area under the curve; KQ=key question; NA=not applicable; NRI=net reclassification improvement; VAL=validity.

© 2018 American Medical Association. All rights reserved.

577 **eTable 41. KQ 2a Fracture prediction studies risk of bias: Part 9, Overall Ratings**

First Author, Year	Justification of Bias Rating	Comments	Overall Judgement of Risk of Bias	Justification of Bias Rating
Ahmed, 2014[51]	Except for perhaps hip fx in women at 5 yrs, calibration plots suggest underestimation of risk at lower risk levels, and overestimation of risk at higher risk levels.	The NRI thresholds used were based on quintiles of the sample distribution of fracture risks. Thresholds used for NRI should be based on sensible and accepted thresholds to define risk groups.	Unclear for AUCs, High for NRIs	NRI risk thresholds not based on sensible/acceptable categories to define risk, they were based on sample distribution. Inadequate followup for 10 year risk prediction.
Azagra, 2011[52]	NR	NR	Unclear	Some concerns about selection bias due to source of study population and attrition of participants over period of followup.
Bauer, 2007[53]	NR	NR	Low	NR
Berry, 2013[54]	NR	NR	Low	NR
Chan, 2012[55]	Low for AUC, High for NRI	The NRI thresholds used were based on tertiles of the sample distribution. Thresholds used for NRI should be based on sensible and accepted thresholds to define risk groups.	Varies by outcome	Unclear for AUC, High For NRI
Chan, 2013[56]	Low for AUC, High for NRI	The NRI thresholds used were based on tertiles of the sample distribution. Thresholds used for NRI should be based on sensible and accepted thresholds to define risk groups.	High	Spectrum bias introducted by subgroup analysis.
Crandall, 2014[57]	NR	NR	Unclear	OST and SCORE were not devleoped and validated to predict fractures; they were developed and validated to predict low BMD/osteoporosis.
Hans, 2011[58]	NR	If multiple DXA scans, just took first one	Low	NR
Hillier, 2007[59]	Removed patients with incident fractures.	NR	Low	NR
Hippisley-Cox, 2012[60]	NR	NR	Unclear	Unclear because of partipant flow

578
579

© 2018 American Medical Association. All rights reserved.

580 **eTable 41. KQ 2a Fracture prediction studies risk of bias: Part 9, Overall Ratings (continued)**

First Author, Year	Justification of Bias Rating	Comments	Overall Judgement of Risk of Bias	Justification of Bias Rating
Iki, 2014[61]	NR	NR	Low	NR
Iki, 2015[62]	Evidence of good calibration	None	Unclear	Length of follow-up only 4.5 years for a 10-year prediction.
Kalvesten, 2016[63]	NR	None	Low	No serious risks of bias. Eligible Outcomes include the discrimination of DXA for predicting fracture, and FRAX (without DXA BMD) for predicting fracture. The diagnostic performance of FRAX for predicting osteoporosis is not eligible because there was 2.1 years between FRAX assessment and DXA measurement. For same reason FRAX w/BMD not eligible as well.
Kanis, 2007[64]	NR	NR	Low	NR
Kwok, 2012[65]	NR	NR	Low	Did not exclude traumatic fractures in definition of "all fractures" but we can just take the data for fragility fractures)
Leslie, 2010[66]	NR	NR	Unclear	Effect of adjustment to final risk category unclear
Leslie, 2012[67]	Model demonstrates the effect of using various non-femoral neck BMD measures	NR	Unclear	Effect of adjustments of absence of data on parental hip fractures unclear
Leslie, 2012[68]	Model demonstrates the effect of not using BMD	NR	Unclear	Effect of adjustments of absence of data on parental hip fractures unclear
Leslie, 2013[69]	NR	NR	Low	NR
Lo, 2011[70]	NR	NR	Unclear	Selection bias and spectrum bias due to how cohort was assembled.
Lundin, 2015[71]	Most of the items are NA.	None	Low	No serious risks of bias

581
582

© 2018 American Medical Association. All rights reserved.

583 **eTable 41. KQ 2a Fracture prediction studies risk of bias: Part 9, Overall Ratings (continued)**

First Author, Year	Justification of Bias Rating	Comments	Overall Judgement of Risk of Bias	Justification of Bias Rating
Melton, 2005[72]	For patients with multiple fractures, only included the first fracture, but unclear if different types of fractures in same person or same types of fracture	NR	High	Due to sampling, definition of outcome
Miller, 2002[73]	NR	NR	High	Self-reported fracture outcomes
Morin, 2009[36]	NR	NR	Unclear	Unclear whether data for OST (age, weight) was collected before fracture for all participants, unclear what proportion of cohort did not have information on predictors
Nguyen, 2004[74]	NR	NR	Unclear	Proportion and management of missing data unclear
Rubin, 2013[75]	NR	NR	Unclear	For short follow-up duration to predict 10 year risk.
Stewart, 2006[76]	NR	NR	Low	NR
van Geel, 2014[77]	NR	NR	Unclear	Follow-up period shorter than instrument

584 **Abbreviations:** AUC= area under the curve; BMD= bone mineral density; DXA=dual energy x-ray absorptiometry; NR=not reported; NRI=net reclassification improvement;
585 OST=osteoporosis self-assessment tool; SCORE=Simple Calculated Osteoporosis Risk Estimation Tool; Yrs=years.

586

© 2018 American Medical Association. All rights reserved.

eTable 42. KQ4 and KQ5 systematic review risk of bias assessments: Part 1, Study Eligibility Criteria

First Author, Year	Describe interventions and comparators (MUST describe usual care and combinations)	Did the review adhere to pre-defined objectives and eligibility criteria?	Were the eligibility criteria appropriate for the review question?	Were eligibility criteria unambiguous?
Crandall et al., 2012[78]	Treatments to prevent fractures vs. Placebo	Yes	Yes	Yes

Abbreviations: KQ=key question; vs= versus.

eTable 43. KQ4 and KQ5 systematic review risk of bias assessments: Part 2, Study Eligibility Criteria

First Author, Year	Were all restrictions in eligibility criteria based on study characteristics appropriate (e.g. date, sample size, study quality, outcomes measured)?	Were any restrictions in eligibility criteria based on sources of information appropriate (e.g. publication status or format, language, availability of data)?	Concerns regarding specification of study eligibility criteria	Did the review search an appropriate range of databases/electronic sources for published and unpublished reports?
Crandall et al., 2012[78]	Yes	Yes	Low	Yes

Abbreviations: KQ=key question.

eTable 44. KQ4 and KQ5 systematic review risk of bias assessments: Part 3, Identification and Selection of Studies

First Author, Year	Were methods additional to database searching used to identify relevant reports?	Were the terms and structure of the search strategy likely to retrieve as many eligible studies as possible?	Were restrictions based on date, publication format, or language appropriate?	Were efforts made to minimize error in selection of studies?
Crandall et al., 2012[78]	Yes	Yes	Probably yes	Yes

Abbreviations: KQ=key question.

eTable 45. KQ4 and KQ5 systematic review risk of bias assessments: Part 4, Identification and Selection of Studies

First Author, Year	Concerns regarding methods used to identify and/or select studies	Were efforts made to minimize error in data collection?	Were sufficient study characteristics available for both review authors and readers to be able to interpret the results?	Were all relevant study results collected for use in the synthesis?
Crandall et al., 2012[78]	Low	No information	Yes	Yes

Abbreviations: KQ=key question.

© 2018 American Medical Association. All rights reserved.

eTable 46. KQ4 and KQ5 systematic review risk of bias assessments: Part 5, Identification and Selection of Studies

First Author, Year	Was risk of bias (or methodological quality) formally assessed using an appropriate tool?	Were efforts made to minimize error in risk of bias assessment?	Concerns regarding methods used to collect data and appraise studies	Did the synthesis include all studies that it should?
Crandall et al., 2012[78]	Yes	No information	Low	Yes

Abbreviations: KQ=key question

eTable 47. KQ4 and KQ5 systematic review risk of bias assessments: Part 6, Synthesis and Findings

First Author, Year	Were all pre-defined analyses reported or departures explained?	Was the synthesis appropriate given the degree of similarity in the research questions, study designs and outcomes across included studies?	Was between-study variation (heterogeneity) minimal or addressed in the synthesis?	Were the findings robust, e.g. as demonstrated through sensitivity analyses?
Crandall et al., 2012[78]	Yes	Yes	Yes	Yes

Abbreviations: KQ=key question

eTable 48. KQ4 and KQ5 systematic review risk of bias assessments: Part 7, Overall Ratings

First Author, Year	Were biases in primary studies minimal or addressed in the synthesis?	Concerns regarding the synthesis	Did the interpretation of findings address all of the concerns identified in Domains 1 to 4?	Was the relevance of identified studies to the review's research question appropriately considered?	Did the reviewers avoid emphasizing results on the basis of their statistical significance?	Risk of bias in the review
Crandall et al., 2012[78]	Yes	Unclear or some concerns	Yes	Yes	Yes	Low

Abbreviations: KQ=key question

© 2018 American Medical Association. All rights reserved.

602 **eTable 49. KQ 4 and 5 risk of bias assessment: Part 1, Participant Selection**

First Author, Year	Describe Interventions and Comparators	Study Design	FOR RCTs: Was Method of Randomization Adequate?	FOR RCTs: Was Allocation Concealment Adequate?	For RCTs: Were There Baseline Imbalances That Suggest A Problem With Randomization?
Abrahamsen, 2010[79]	G1: Alendronate G2: Untreated	Cohort study	NA-not an RCT	NA-not an RCT	NA-not an RCT
Adachi, 2009[80]	G1: alendronate 10 mg daily (generic preparation) G2: Placebo	RCT parallel	Yes	Yes	Probably yes
Barrett-Connor, 2002[81]	G1: Raloxifene (60mg/day) G2: Raloxifene (120mg/day) G2: Placebo	Post-hoc or subgroup analysis of RCT	Yes	Yes	No
Barrett-Connor, 2004[82]	G1: Raloxifene (60mg/day or 120mg/day) G2: Placebo	RCT parallel	Yes	Yes	No
Bone, 2000[83]	G1: Alendonate 10 mg /day G2: conjugage equine estrogen 0.625 mg /day) G3: Alendronate + CEE G4: placebo	RCT parallel	Yes	No information	No
Bone, 2008[84]	G1: Denosumab G2: Placebo	RCT parallel	Probably yes	Probably yes	No
Boonen, 2012[85]	G1: intravenous infusion of zoledronic acid (5 mg) for 12 months G2: Placebo	RCT parallel	Yes	yes	No
Cartsos, 2008[86]	Intervention: Bisphosphonate use Comparator: no bisphosphonate use	Case-control (how they described)	NA-not an RCT	NA-not an RCT	NA-not an RCT
Chapurlat, 2013[87]	G1: ibandronate G2: placebo	RCT parallel	Probably yes	Yes	No
Cryer, 2005[88]	G1: alendronate 70 mg weekly G2: placebo	RCT parallel	Yes	Yes	No
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	G1: alendronate 5mg per day for 2 years, then 10 mg per day for 3 years G2: placebo	RCT parallel	Yes	Yes	No
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	G1: Denosumab G2: Placebo	RCT parallel	Probably yes	Probably yes	No

603
604

© 2018 American Medical Association. All rights reserved.

605 **eTable 49. KQ 4 and 5 risk of bias assessment: Part 1, Participant Selection (continued)**

First Author, Year	Describe Interventions and Comparators	Study Design	FOR RCTs: Was Method of Randomization Adequate?	FOR RCTs: Was Allocation Concealment Adequate?	For RCTs: Were There Baseline Imbalances That Suggest A Problem With Randomization?
Eisman, 2004[96]	G1:alendronate 70 mg weekly G2: placebo	RCT parallel	Yes	Yes	No
Fogelman, 2000[97]	G1: Risedronate 5 mg/d X 24 months G2: Placebo	RCT parallel	No information	No information	No
Greenspan, 2002[98]	G1: alendronate 70 mg weekly G2: placebo	RCT parallel	No information	No information	No
Greenspan, 2003[99]	G1: Alendonate 10 mg /day G2: conjugated equine estrogen 0.625 mg /day with or without medroxyprogesterone 2.5mg daily based on uterus presence G3: Alendronate + CEE G4: placebo	RCT parallel	Yes	Yes	No
Grey, 2010[100]	G1: Zolendronate 5 mg IV x 1 dose G2: Placebo	RCT parallel	Yes	Yes	Probably yes
Hosking, 2003[101]	G1: Risedronate 5 mg/d X 3 months G2: Alendronate 70 mg/once weekly X 3 months G3: Placebo	RCT parallel	Yes	Yes	No
Hosking, 2003[101]	G1: alendronate 70 mg weekly G2: Risendronate 5 mg daily G3: Placebo	RCT parallel	Yes	Yes	No
Johnell, 2002[102]	RLX 60, placebo	RCT parallel	Yes	Yes	Probably no
Keech, 2005[103]	G1: Raloxifene 60 mg / day G2 : Placebo	Post-hoc or subgroup analysis of RCT	Yes	Yes	No
Kung, 2000[104]	G1: alendronate 10 mg daily G2: placebo	RCT parallel	No information	No information	No
Lasco, 2011[105]	G1: Teriparatide + calcium + vitamin D G2: Calcium + vitamin D	Cohort study	NA-not an RCT	NA-not an RCT	NA-not an RCT
Lewiecki, 2007[106]	G1: Denosumab (included varying dosages over 3 and 6 months) G2: Alenbronate G3: Placebo	RCT parallel	Probably yes	Probably yes	No
McCloskey, 2012[107]	G1: 60 mg Denosumab SC q 6 mos for 36 mos G2: Placebo	RCT parallel	No information	No information	No

606

© 2018 American Medical Association. All rights reserved.

607 **eTable 49. KQ 4 and 5 risk of bias assessment: Part 1, Participant Selection (continued)**

First Author, Year	Describe Interventions and Comparators	Study Design	FOR RCTs: Was Method of Randomization Adequate?	FOR RCTs: Was Allocation Concealment Adequate?	For RCTs: Were There Baseline Imbalances That Suggest A Problem With Randomization?
McClung, 2004[108]	G1: 0.5mg ibandronate daily G2: 1.0mg ibandronate daily G3: 2.5mg ibandronate daily G4:placebo	RCT parallel	No information	No information	No
McClung, 2006[109]	G1: Lasofoxifene 0.25 mg/day G2: Lasofoxifene 1.0 mg/day G3: Raloxifene 60 mg/day G4: Placebo	RCT parallel	No information	No information	No
McClung, 2006[110]	G1: Denosumab 6 mg Q3mo G2: Denosumab 14 mg Q3mo G3: Denosumab 30 mg Q3mo G4: Denosumab 14 mg Q6mo G5: Denosumab 60 mg Q6mo G6: Denosumab 100 mg Q6mo G7: Denosumab 210 mg Q6mo G8: Alendronate 70 mg weekly G9: placebo	RCT parallel	No information	No information	No
McClung, 2009[111]	G1: Zoledronic acid 5 mg IV q 12 mos for 2 doses G2: Zoledronic acic 5mg IV once and placebo at 12 mos G3: Placebo at baseline and 12 mos	RCT parallel	Yes	Yes	No
Meunier, 1999[112]	raloxifene, 60 mg, 150 mg or placebo	RCT parallel	No information	Probably yes	No
Miller, 2008[113]	G1: Bazedoxifene 10 mg G2: Bazedoxifene 20 mg G3: Bazedoxifene 40 mg G4: Raloxifene 60 mg G5: Placebo	RCT parallel	yes	yes	No
Morii, 2003[114]	raloxifene, 2 dosage amounts vs placebo	RCT parallel	No information	No information	Probably no
Murphy, 2001[115]	G1: MK-677 25 mg daily G2: alendronate 10 mg daily G3: MK-677 and alendronate G4: placebo **pull out G2 and G4 data only for KQ5	RCT parallel	Yes	Yes	No
Nakamura, 2012[116]	G1: Denosumab 14 mg G2: Denosumab 60 mg G3: Denosumab 100 mg G4: Placebo	RCT parallel	No information	No information	No

608

© 2018 American Medical Association. All rights reserved.

609 **eTable 49. KQ 4 and 5 risk of bias assessment: Part 1, Participant Selection (continued)**

First Author, Year	Describe Interventions and Comparators	Study Design	FOR RCTs: Was Method of Randomization Adequate?	FOR RCTs: Was Allocation Concealment Adequate?	For RCTs: Were There Baseline Imbalances That Suggest A Problem With Randomization?
Orwoll, 2003[117]	G1: 20 µg teriparatide: 151 G2:40 µg teriparatide: 139 G3: placebo: 147	RCT parallel	Yes	Yes	No
Pazianas, 2008[118]	Intervention: oral bisphosphate use Comparator: No bisphosphate use	case-control (how they described)	NA-not an RCT	NA-not an RCT	NA-not an RCT
Ravn, 1996[119]	G1: 0.25mg ibandronate daily G2: 0.5mg ibandronate daily G3: 1.0mg ibandronate daily G4: 2.5mg ibandronate daily G5: 5.0 mg ibandronate daily G6: placebo	RCT parallel	No information	No information	No
Reginster, 2005[120]	G1: 50mg ibandronate monthly 1 month, followed by 50 mg monthly 2 months for half the sample and 100 mg monthly for 2 months for the other half G2: 100mg ibandronate monthly for 3 months G3: 150mg ibandronate monthly for 3 months G4: placebo for 3 months	RCT parallel	No information	No information	Yes
Rhee, 2012[121]	G1: Bisphosphonate use G2: non bisphosphonate use	Cohort study	NA-not an RCT	NA-not an RCT	NA-not an RCT
Riis, 2001[122]	G1: 2.5mg ibandronate daily continuous therapy G2: 20mg ibandronate intermittent cyclical therapy G3: Placebo	RCT parallel	No information	No information	No
Samelson, 2014[123]	G1: 60 mg Denosumab SC q 6 mos for 36 mos G2: Placebo	Post-hoc or subgroup analysis of RCT	No information	No information	Probably no
Shiraki, 2003[124]	G1: Risedronate 5 mg/d X 36 weeks G2: Placebo	RCT cluster	No information	No information	No
Simon, 2013[125]	G1: 60 mg Denosumab SC q 6 mos for 36 mos G2: Placebo	RCT parallel	No information	No information	No
Sontag, 2010[126]	G1: Raloxifene in women without baseline prevalent vertebral fracture G2: Placebo in women without baseline prevalent vertebral fracture	Post-hoc or subgroup analysis of RCT	Yes	Yes	No

610

© 2018 American Medical Association. All rights reserved.

eTable 49. KQ 4 and 5 risk of bias assessment: Part 1, Participant Selection (continued)

First Author, Year	Describe Interventions and Comparators	Study Design	FOR RCTs: Was Method of Randomization Adequate?	FOR RCTs: Was Allocation Concealment Adequate?	For RCTs: Were There Baseline Imbalances That Suggest A Problem With Randomization?
Sorensen, 2008[127]	G1: bisphosphonate therapy* G2: placebo *Study examined all busphosphonates used in Danish prescription database, predominently alendronate, etidronate. Only 5 control patients used risendronate. No patients used zolendronic acid.	Case-control	NA-not an RCT	NA-not an RCT	NA-not an RCT
Tanko, 2003[128]	G1: 5mg ibandronate weekly G2: 10mg ibandronate weekly G3: 20mg ibandronate weekly G4: placebo	RCT parallel	No information	No information	No
Thiebaud, 1997[129]	G1: 2.5mg ibandronate IV every 3 months G2: .5mg ibandronate IV every 3 months G3: 1mg ibandronate IV every 3 months G4: 2mg ibandronate IV every 3 months G5: placebo	RCT parallel	No information	No information	No
Tucci, 1996[130]	G1: Alendronate 5mg daily G2: Alendronate 10 mg daily G3: Alendronate 20 daily for 2 years then 5 mg daily for 1 year G4: placebo	RCT parallel	Yes	Yes	No
Van Staa, 1997[131]	G1: Cyclinical Etidronate (1 or more cyclical etidronate prescriptions) G2: Nonosteoporosis control (as recorded in their medical records and no bisphosphonate use)	Cohort study	NA-not an RCT	NA-not an RCT	NA-not an RCT
Vestergaard, 2010[132]	Gastric & esophagus events	Cohort study	NA-not an RCT	NA-not an RCT	NA-not an RCT
Vestergaard, 2011[133]	Stroke	Cohort study	NA-not an RCT	NA-not an RCT	NA-not an RCT
Vestergaard, 2012[134]	Cardiac and atherosclerosis	Cohort study	NA-not an RCT	NA-not an RCT	NA-not an RCT
Vestergaard, 2011[135]	Femoral shaft and subtrochanteric fractures	Cohort study	NA-not an RCT	NA-not an RCT	NA-not an RCT
Vestergaard, 2012[136]	Jaw disease	Cohort study	NA-not an RCT	NA-not an RCT	NA-not an RCT

Abbreviations: CEE=conjugated equine estrogen;G=group; KQ=key question; mg=milligram; mg/d=milligram per day; mo=month; NA=not applicable; RCT=randomized controlled trials.

© 2018 American Medical Association. All rights reserved.

615 **eTable 50. KQ 4 and 5 risk of bias assessment: Part 2, Participant Selection**

First Author, Year	FOR COHORTs: Was Selection into the Study Unrelated to Intervention or Unrelated to Outcome?	FOR COHORTs: Do Start of Follow-Up and Start of Intervention Coincide for Most Participants?	FOR COHORTs: Were Adjustment Techniques Used That are Likely to Correct for the Presence of Selection Biases?	FOR CASE-CONTROLS: Were the Controls Sampled from the Population that Gave Rise to the Cases, or Using Another Method That Avoids Selection Bias?	Bias Arising from Randomization or Selection?	Comments
Abrahamsen, 2010[79]	Probably no	Yes	Yes	NA-not a case-control	Probably no	Women treated with alendronate by definition have increased risk of fracture, prompting their treatment with the drug.
Adachi, 2009[80]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably yes	Alendronate group had greater proportion of patients with history of UGI disease, active UGI disease, esophogeal disease, no statistical comparison is given, but the differences are large enough to warrant some concern for risk of bias as it does not appear that these differences were corrected for in the analysis.
Barrett-Connor, 2002[81]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Barrett-Connor, 2004[82]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	Not enough information on randomization process.
Bone, 2000[83]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Bone, 2008[84]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	No information on allocation concealment
Boonen, 2012[85]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Cartsos, 2008 [86]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	Yes	data comes from medical claims data; possible errors in coding; does not include uninsured; sample not representative of total population
Chapurlat, 2013[87]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Cryer, 2005[88]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR

616
617

© 2018 American Medical Association. All rights reserved.

618 **eTable 50. KQ 4 and 5 risk of bias assessment: Part 2, Participant Selection (continued)**

First Author, Year	FOR COHORTs: Was Selection into the Study Unrelated to Intervention or Unrelated to Outcome?	FOR COHORTs: Do Start of Follow-Up and Start of Intervention Coincide for Most Participants?	FOR COHORTs: Were Adjustment Techniques Used That are Likely to Correct for the Presence of Selection Biases?	FOR CASE-CONTROLS: Were the Controls Sampled from the Population that Gave Rise to the Cases, or Using Another Method That Avoids Selection Bias?	Bias Arising from Randomization or Selection?	Comments
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	No information on allocation concealment
Eisman, 2004[96]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Fogelman, 2000[97]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	NR
Greenspan, 2002[98]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	The article does not include information on randomization or concealment
Greenspan, 2003[99]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Grey, 2010[100]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably yes	The authors did not clearly adjust for baseline fracture.
Hosking, 2003[101]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Hosking, 2003[101]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Johnell, 2002[102]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Keech, 2005[103]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Kung, 2000[104]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	The article does not include information on randomization or concealment

619

© 2018 American Medical Association. All rights reserved.

620 **eTable 50. KQ 4 and 5 risk of bias assessment: Part 2, Participant Selection (continued)**

First Author, Year	FOR COHORTs: Was Selection into the Study Unrelated to Intervention or Unrelated to Outcome?	FOR COHORTs: Do Start of Follow-Up and Start of Intervention Coincide for Most Participants?	FOR COHORTs: Were Adjustment Techniques Used That are Likely to Correct for the Presence of Selection Biases?	FOR CASE-CONTROLS: Were the Controls Sampled from the Population that Gave Rise to the Cases, or Using Another Method That Avoids Selection Bias?	Bias Arising from Randomization or Selection?	Comments
Lasco, 2011[105]	No	Yes	NA	NA-not a case-control	Yes	One arm has osteoporosis and other has osteopenia; the differences between arms could have served as a prognostic factor and contribute to confounding.
Lewiecki, 2007[106]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	No information on allocation concealment.
McCloskey, 2012[107]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No information	NR
McClung, 2004[108]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably no	no information provided on method of randomization or concealment
McClung, 2006[109]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	Not enough information on randomization process
McClung, 2006[110]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	NR
McClung, 2009[111]	No	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	NR
Meunier, 1999[112]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	NR
Miller, 2008[113]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	NR
Morii, 2003[114]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	Some missing info
Murphy, 2001[115]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Nakamura, 2012[116]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	No information provided on method of randomization or concealment
Orwoll, 2003[117]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Pazianas, 2008[118]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	Yes	data comes from medical claims data; possible errors in coding; does not include uninsured; sample not representative of total population
Ravn, 1996[119]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably no	No information provided on method of randomization or concealment

621
622

© 2018 American Medical Association. All rights reserved.

623 **eTable 50. KQ 4 and 5 risk of bias assessment: Part 2, Participant Selection (continued)**

First Author, Year	FOR COHORTs: Was Selection into the Study Unrelated to Intervention or Unrelated to Outcome?	FOR COHORTs: Do Start of Follow-Up and Start of Intervention Coincide for Most Participants?	FOR COHORTs: Were Adjustment Techniques Used That are Likely to Correct for the Presence of Selection Biases?	FOR CASE-CONTROLS: Were the Controls Sampled from the Population that Gave Rise to the Cases, or Using Another Method That Avoids Selection Bias?	Bias Arising from Randomization or Selection?	Comments
Reginster, 2005[120]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably no	Absence of specific BMD criteria led to the inclusion of some participants were not osteoporotic
Rhee, 2012[121]	Yes	No	No information	NA-not a case-control	Probably yes	Although the authors attempt to create an new user cohort by excluded anyone with a prescription for 16 months prior to the observation of the outcome, it's unclear whether and how many participants might have been exposed to osteoporosis drugs before that period and stopped taking medications.
Riis, 2001[122]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably no	No information provided on method of randomization or concealment
Samelson, 2014[123]	Yes	Yes	No information	NA-not a case-control	Probably no	NR
Shiraki, 2003[124]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	Probably no	NR
Simon, 2013[125]	Yes	Yes	Probably yes	NA-not a case-control	Probably no	No detail on method of randomization and allocation concealment.
Sontag, 2010[126]	Yes	Yes	NA	NA-not a case-control	Probably no	NR
Sorensen, 2008[127]	NA-not a cohort	NA-not a cohort	NA-not a cohort	Yes	No	NR
Tanko, 2003[128]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably no	no information provided on method of randomization or concealment
Thiebaud, 1997[129]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably no	no information provided on method of randomization or concealment Slight differences length of menopause
Tucci, 1996[130]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a case-control	No	NR
Van Staa, 1997[131]	Yes	Yes	Yes	NA-not a case-control	Probably no	NR

624
625

© 2018 American Medical Association. All rights reserved.

626 **eTable 50. KQ 4 and 5 risk of bias assessment: Part 2, Participant Selection (continued)**

First Author, Year	FOR COHORTs: Was Selection into the Study Unrelated to Intervention or Unrelated to Outcome?	FOR COHORTs: Do Start of Follow-Up and Start of Intervention Coincide for Most Participants?	FOR COHORTs: Were Adjustment Techniques Used That are Likely to Correct for the Presence of Selection Biases?	FOR CASE-CONTROLS: Were the Controls Sampled from the Population that Gave Rise to the Cases, or Using Another Method That Avoids Selection Bias?	Bias Arising from Randomization or Selection?	Comments
Vestergaard, 2010[132]	Yes	Yes	Irrelevant, claim there is no missing data	NA-not a case-control	No	NR
Vestergaard, 2011[133]	Yes	Yes	Irrelevant, claim there is no missing data	Yes	No	NR
Vestergaard, 2012[134]	Yes	Yes	Irrelevant, claim there is no missing data	NA-not a case-control	No	NR
Vestergaard, 2011[135]	Yes	Yes	Irrelevant, claim there is no missing data	NA-not a case-control	No	NR
Vestergaard, 2012[136]	Yes	Yes	Irrelevant, claim there is no missing data	NA-not a case-control	No	NR

627 **Abbreviations:** NA=not applicable; NR=not reported.

628

© 2018 American Medical Association. All rights reserved.

eTable 51. KQ 4 and 5 risk of bias assessment: Part 3, Confounding

First Author, Year	FOR COHORTS AND CASE CONTROLS: Is Confounding of the Effect of Intervention Unlikely in this Study?	FOR COHORTS: Were Participants Analyzed According to Their Initial Intervention Group Throughout Follow Up?	FOR COHORT STUDIES: Were Intervention Discontinuations or Switches Unlikely to be Related to Factors That are Prognostic for the Outcome?	FOR COHORT AND CASE-CONTROL STUDIES: Did the Authors use an Appropriate Analysis Method That Adjusted for all the Critically Important Confounding Domains?
Abrahamsen, 2010[79]	Probably no	Yes	No information	Probably yes
Adachi, 2009[80]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Barrett-Connor, 2002[81]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Barrett-Connor, 2004[82]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Bone, 2000[83]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Bone, 2008[84]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Boonen, 2012[85]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Cartsos, 2008 [86]	Probably no	NA-not a cohort	NA-not a cohort	No information
Chapurlat, 2013[87]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Cryer, 2005[88]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Eisman, 2004[96]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Fogelman, 2000[97]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Greenspan, 2002[98]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Greenspan, 2003[99]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Grey, 2010[100]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Hosking, 2003[101]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Hosking, 2003[101]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Johnell, 2002[102]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Keech, 2005[103]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Kung, 2000[104]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Lasco, 2011[105]	No	Yes	No information	No information
Lewiecki, 2007[106]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
McCloskey, 2012[107]	Probably yes	Yes	Yes	Probably yes
McClung, 2004[108]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
McClung, 2006[109]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort

© 2018 American Medical Association. All rights reserved.

632 **eTable 51. KQ 4 and 5 risk of bias assessment: Part 3, Confounding (continued)**

First Author, Year	FOR COHORTS AND CASE CONTROLS: Is Confounding of the Effect of Intervention Unlikely in this Study?	FOR COHORTS: Were Participants Analyzed According to Their Initial Intervention Group Throughout Follow Up?	FOR COHORT STUDIES: Were Intervention Discontinuations or Switches Unlikely to be Related to Factors That are Prognostic for the Outcome?	FOR COHORT AND CASE-CONTROL STUDIES: Did the Authors use an Appropriate Analysis Method That Adjusted for all the Critically Important Confounding Domains?
McClung, 2006[110]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
McClung, 2009[111]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Meunier, 1999[112]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Miller, 2008[113]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Morii, 2003[114]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Murphy, 2001[115]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Nakamura, 2012[116]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Orwoll, 2003[117]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Pazianas, 2008 [118]	Probably no	NA-not a cohort	NA-not a cohort	Yes
Ravn, 1996[119]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Reginster, 2005[120]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Rhee, 2012[121]	Yes	Yes	Unclear, all switches dropped from analysis	NA
Riis, 2001[122]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Samelson, 2014[123]	Probably yes	Yes	Yes	Probably yes
Shiraki, 2003[124]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Simon, 2013[125]	Probably yes	Yes	Yes	Probably yes
Sontag, 2010[126]	Yes	NA	Yes	No
Sorensen, 2008[127]	No	NA-not a cohort	NA-not a cohort	Yes
Tanko, 2003[128]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Thiebaud, 1997[129]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Tucci, 1996[130]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort
Van Staa, 1997[131]	Yes	NA	NA	NA
Vestergaard, 2010[132]	No	No information	No information	No
Vestergaard, 2011[133]	No	No information	No information	No
Vestergaard, 2012[134]	No	No information	No information	No
Vestergaard, 2011[135]	No information	No information	No	Yes
Vestergaard, 2012[136]	No	No information	No information	No

633 **Abbreviations:** KQ=key question; NA=not applicable.

634
635

© 2018 American Medical Association. All rights reserved.

636 **eTable 52. KQ 4 and 5 risk of bias assessment: Part 4, Confounding**

First Author, Year	FOR COHORT STUDIES: Did the Authors Avoid Adjusting for Post Intervention Variables?	FOR COHORT STUDIES Did the Authors use an Appropriate Analysis Method That Adjusted for all the Critically Important Confounding Domains and for Time-Varying Confounding?	Bias Arising from Confounding?	Comments
Abrahamsen, 2010[79]	Yes	Probably yes	Probably yes	NR
Adachi, 2009[80]	NA-not a cohort	NA-not a cohort	No	NR
Barrett-Connor, 2002[81]	NA-not a cohort	NA-not a cohort	NA	NR
Barrett-Connor, 2004[82]	NA-not a cohort	NA-not a cohort	NA	NR
Bone, 2000[83]	NA-not a cohort	NA-not a cohort	No	NR
Bone, 2008[84]	NA-not a cohort	NA-not a cohort	no	NR
Boonen, 2012[85]	NA-not a cohort	NA-not a cohort	No	NR
Cartsos, 2008 [86]	NA-not a cohort	NA-not a cohort	Probably yes	Possible patients could have been taking other treatments that were not documented; no mention of how confounding was handled or if considered
Chapurlat, 2013[87]	NA-not a cohort	NA-not a cohort	N/A	NR
Cryer, 2005[88]	NA-not a cohort	NA-not a cohort	No	NR
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	NA-not a cohort	NA-not a cohort	No	NR
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	NA-not a cohort	NA-not a cohort	No	NR
Eisman, 2004[96]	NA-not a cohort	NA-not a cohort	No	NR
Fogelman, 2000[97]	NA-not a cohort	NA-not a cohort	No information	NR
Greenspan, 2002[98]	NA-not a cohort	NA-not a cohort	No	NR
Greenspan, 2003[99]	NA-not a cohort	NA-not a cohort	No	NR
Grey, 2010[100]	NA-not a cohort	NA-not a cohort	No	NR
Hosking, 2003[101]	NA-not a cohort	NA-not a cohort	No information	NR
Hosking, 2003[101]	NA-not a cohort	NA-not a cohort	No	NR
Johnell, 2002[102]	NA-not a cohort	NA-not a cohort	Probably no	NR
Keech, 2005[103]	NA-not a cohort	NA-not a cohort	NA	NR
Kung, 2000[104]	NA-not a cohort	NA-not a cohort	No	NR

637
638

© 2018 American Medical Association. All rights reserved.

639 **eTable 52. KQ 4 and 5 risk of bias assessment: Part 4, Confounding (continued)**

First Author, Year	FOR COHORT STUDIES: Did the Authors Avoid Adjusting for Post Intervention Variables?	FOR COHORT STUDIES Did the Authors use an Appropriate Analysis Method That Adjusted for all the Critically Important Confounding Domains and for Time-Varying Confounding?	Bias Arising from Confounding?	Comments
Lasco, 2011[105]	No information	No information	Yes	One arm has osteoporosis and other has osteopenia; the differences between arms could have served as a prognostic factor and contribute to confounding
Lewiecki, 2007[106]	NA-not a cohort	NA-not a cohort	No	NR
McCloskey, 2012[107]	Probably yes	NA	No information	The analysis was prespecified according to the methods and does not appear to be a subgroup. The are looking at efficacy across the range of baseline FRAX risk.
McClung, 2004[108]	NA-not a cohort	NA-not a cohort	NA	NA, Not a cohort or case control
McClung, 2006[109]	NA-not a cohort	NA-not a cohort	No	Not a cohort or case control
McClung, 2006[110]	NA-not a cohort	NA-not a cohort	No	NR
McClung, 2009[111]	NA-not a cohort	NA-not a cohort	No	RCT design mitigates risk of confounding from known and unknown factors.
Meunier, 1999[112]	NA-not a cohort	NA-not a cohort	Probably no	NR
Miller, 2008[113]	NA-not a cohort	NA-not a cohort	No	NR
Morii, 2003[114]	NA-not a cohort	NA-not a cohort	Probably no	NR
Murphy, 2001[115]	NA-not a cohort	NA-not a cohort	No	NR
Nakamura, 2012[116]	NA-not a cohort	NA-not a cohort	NA	NR
Orwoll, 2003[117]	NA-not a cohort	NA-not a cohort	No	Not a cohort study
Pazianas, 2008 [118]	NA-not a cohort	NA-not a cohort	Probably no	Possible patients could have been taking other treatments that were not documented
Ravn, 1996[119]	NA-not a cohort	NA-not a cohort	NA, Not a cohort or case control	NR
Reginster, 2005[120]	NA-not a cohort	NA-not a cohort	NA, Not a cohort or case control	NR
Rhee, 2012[121]	No	No	Yes	They also dropped all patients with switches, which potentially selectively drops patients with reactions to initial drug therapy

© 2018 American Medical Association. All rights reserved.

640 **eTable 52. KQ 4 and 5 risk of bias assessment: Part 4, Confounding (continued)**

First Author, Year	FOR COHORT STUDIES: Did the Authors Avoid Adjusting for Post Intervention Variables?	FOR COHORT STUDIES Did the Authors use an Appropriate Analysis Method That Adjusted for all the Critically Important Confounding Domains and for Time-Varying Confounding?	Bias Arising from Confounding?	Comments
Riis, 2001[122]	NA-not a cohort	NA-not a cohort	NA, Not a cohort or case control	NR
Samelson, 2014[123]	Yes	NA, if item 10 is yes/probably yes	Probably no	Treatment assignment is random; CV risks were balanced between groups.
Shiraki, 2003[124]	NA-not a cohort	NA-not a cohort	No information	NR
Simon, 2013[125]	Probably yes	NA	Probably no	NR
Sontag, 2010[126]	No	NA	Yes	During a 1-year extension phase, women were permitted to take other bone-active agents, except for oral estrogen or estrogen–progestin therapy. 16.4% and 12.3% of women in the placebo and raloxifene 60 mg/day groups, respectively, used other bone-active a
Sorensen, 2008[127]	NA-not a cohort	NA-not a cohort	No	NR
Tanko, 2003[128]	NA-not a cohort	NA-not a cohort	NA, Not a cohort or case control	NR
Thiebaud, 1997[129]	NA-not a cohort	NA-not a cohort	NA, Not a cohort or case control	NR
Tucci, 1996[130]	NA-not a cohort	NA-not a cohort	No	NR
Van Staa, 1997[131]	NA	NA	No	NR
Vestergaard, 2010[132]	Yes	Probably no	Probably yes	Given the results it's likely that there were other underlying variables that they didn't fully account for. For example, are all NSAIDS in the drugs registry? What about OTC NSAIDS? Given that a third of their sample had fractures, likely they had pain t
Vestergaard, 2011[133]	Yes	Probably no	Probably yes	NR

641
642

© 2018 American Medical Association. All rights reserved.

643 **eTable 52. KQ 4 and 5 risk of bias assessment: Part 4, Confounding (continued)**

First Author, Year	FOR COHORT STUDIES: Did the Authors Avoid Adjusting for Post Intervention Variables?	FOR COHORT STUDIES Did the Authors use an Appropriate Analysis Method That Adjusted for all the Critically Important Confounding Domains and for Time-Varying Confounding?	Bias Arising from Confounding?	Comments
Vestergaard, 2012[134]	Yes	Probably no	Probably yes	Given the results it's likely that there were other underlying variables that they didn't fully account for. For example, did they fully control for all other causes of MI such as smoking and hypertension.
Vestergaard, 2011[135]	Probably no	Probably yes	No	NR
Vestergaard, 2012[136]	Yes	Probably no	Probably yes	NR

644 **Abbreviations:** FRAX=Fracture Risk Assessment tool; KQ=key question; MI=myocardial infarction; NA=not applicable; NR=not reported; NSAIDS=nonsteroidal anti-
645 inflammatory drugs; OTC=over the counter.

646

© 2018 American Medical Association. All rights reserved.

647 **eTable 53. KQ 4 and 5 risk of bias assessment: Part 5, Intervention Measurement**

First Author, Year	FOR COHORTS AND CASE CONTROLS: Is Intervention Status Well Defined?	FOR COHORTS AND CASE CONTROLS: Was Information on Intervention Status Recorded at the Time of Intervention?	FOR COHORTS AND CASE CONTROLS Was Information on Intervention Status Unaffected by Knowledge of the Outcome or Risk of the Outcome?	Bias Arising from Measurement of the Intervention?	Comments
Abrahamsen, 2010[79]	Yes	Yes	Yes	Probably no	NR
Adachi, 2009[80]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Barrett-Connor, 2002[81]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA	NR
Barrett-Connor, 2004[82]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Bone, 2000[83]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Bone, 2008[84]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Boonen, 2012[85]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA-not a cohort	NR
Cartsos, 2008[86]	No	No	Probably yes	Yes	Intervention based on dispensing information from claims data, information on dose not available
Chapurlat, 2013[87]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA	NR
Cryer, 2005[88]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Eisman, 2004[96]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Fogelman, 2000[97]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No information	NR
Greenspan, 2002[98]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Greenspan, 2003[99]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Grey, 2010[100]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Hosking, 2003[101]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No information	NR
Hosking, 2003[101]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Johnell, 2002[102]	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably no	NR
Keech, 2005[103]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA	NR
Kung, 2000[104]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Lasco, 2011[105]	Yes	Yes	No information	Probably no	NR
Lewiecki, 2007[106]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
McCloskey, 2012[107]	Yes	Yes	Yes	No	It was prespecified.

648

© 2018 American Medical Association. All rights reserved.

649 **eTable 53. KQ 4 and 5 risk of bias assessment: Part 5, Intervention Measurement (continued)**

First Author, Year	FOR COHORTS AND CASE CONTROLS: Is Intervention Status Well Defined?	FOR COHORTS AND CASE CONTROLS: Was Information on Intervention Status Recorded at the Time of Intervention?	FOR COHORTS AND CASE CONTROLS Was Information on Intervention Status Unaffected by Knowledge of the Outcome or Risk of the Outcome?	Bias Arising from Measurement of the Intervention?	Comments
McClung, 2004[108]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA, not a cohort or case control	NR
McClung, 2006[109]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	not a cohort or case control
McClung, 2006[110]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
McClung, 2009[111]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	RCT Design so all items NA.
Meunier, 1999[112]	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably no	NR
Miller, 2008[113]	NA-not a cohort	NA-not a cohort	NA-not a cohort	no	NR
Morii, 2003[114]	NA-not a cohort	NA-not a cohort	NA-not a cohort	Probably no	NR
Murphy, 2001[115]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Nakamura, 2012[116]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA	NR
Orwoll, 2003[117]	NA-not a cohort	NA-not a cohort	NA-not a cohort	no	NR
Pazianas, 2008[118]	No	No	Probably yes	Yes	Intervention based on dispensing information from claims data, information on dose etc. Not available
Ravn, 1996[119]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA, not a cohort or case control	NR
Reginster, 2005[120]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA, not a cohort or case control	NR
Rhee, 2012[121]	Yes	Yes	Yes	No	NR
Riis, 2001[122]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA, not a cohort or case control	NR
Samelson, 2014[123]	Probably yes	Yes	Yes	Probably no	NR
Shiraki, 2003[124]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No information	NR
Simon, 2013[125]	Yes	Yes	Yes	Probably no	NR
Sontag, 2010[126]	Yes	Yes	Yes	Probably no	NR
Sorensen, 2008[127]	Yes	Yes	Yes	No	NR
Tanko, 2003[128]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA, not a cohort or case control	NR
Thiebaud, 1997[129]	NA-not a cohort	NA-not a cohort	NA-not a cohort	NA, not a cohort or case control	NR

650

© 2018 American Medical Association. All rights reserved.

651 **eTable 53. KQ 4 and 5 risk of bias assessment: Part 5, Intervention Measurement (continued)**

First Author, Year	FOR COHORTS AND CASE CONTROLS: Is Intervention Status Well Defined?	FOR COHORTS AND CASE CONTROLS: Was Information on Intervention Status Recorded at the Time of Intervention?	FOR COHORTS AND CASE CONTROLS Was Information on Intervention Status Unaffected by Knowledge of the Outcome or Risk of the Outcome?	Bias Arising from Measurement of the Intervention?	Comments
Tucci, 1996[130]	NA-not a cohort	NA-not a cohort	NA-not a cohort	No	NR
Van Staa, 1997[131]	No	Yes	Yes	Yes	Intervention status defined as patients who had received a prescription
Vestergaard, 2010[132]	No	No information	Yes	Probably yes	NR
Vestergaard, 2011[133]	No	No information	Yes	Probably yes	NR
Vestergaard, 2012[134]	No	No information	Yes	Probably yes	NR
Vestergaard, 2011[135]	Yes	Probably yes	None	No	NA, no attrition
Vestergaard, 2012[136]	No	No information	Yes	Probably yes	NR

652 **Abbreviations:** NA=not applicable; NR=not reported; RCT=randomized controlled trials.

653
654

© 2018 American Medical Association. All rights reserved.

655 **eTable 54. KQ 4 and 5 risk of bias assessment: Part 6, Missing Data**

First Author, Year	FOR RCTS and COHORTS: What was the overall attrition? What was the attrition by group? Did attrition vary for different outcomes?	FOR RCTS and COHORTS: Did the study have high attrition raising concern for bias?	FOR RCTS and COHORTS: Are the proportion of participants and reasons for missing data similar across interventions?	FOR CASE-CONTROLS: Are the proportion of participants and reasons for missing data similar across cases and controls?
Abrahamsen, 2010[79]	Overall: NR G1: 3.1% G2: 3.0% Vary by outcome: Probably no	No	Yes	NA-not a case-control
Adachi, 2009[80]	Overall:16.2 [%] G1: 18.6 [%] G2: 11.6% [%] Vary by Outcome? No	No	Yes	NA-not a case-control
Barrett-Connor, 2002[81]	Overall:26% G1: 26% G2: 25% G3: 26% Vary by Outcome? No	Yes	No	NA
Barrett-Connor, 2004[82]	Overall: 26% G1: 26.2 G2: 25.2 G3: 26.4 Vary by Outcome? no	No	Yes	NA
Bone, 2000[83]	Overall: 24.7 [%] G1: 24/92=26% G4: 16/50=32% Other reasons for attrition: withdrew conset, lost to follow-up, protocol violations, no signficant varation between groups	Yes	Yes	NA-not a case-control
Bone, 2008[84]	Overall attrition: 3/332=0.09% G1: 2/166 (1.2%) G2: 1/166 (0.06%)	No	Yes	NA-not a cohort
Boonen, 2012[85]	Overall: 11% G1: 10% G2: 12% Vary by Outcome? No	No	Yes	NA-not a case control

656

© 2018 American Medical Association. All rights reserved.

eTable 54. KQ 4 and 5 risk of bias assessment: Part 6, Missing Data (continued)

First Author, Year	FOR RCTS and COHORTS: What was the overall attrition? What was the attrition by group? Did attrition vary for different outcomes?	FOR RCTS and COHORTS: Did the study have high attrition raising concern for bias?	FOR RCTS and COHORTS: Are the proportion of participants and reasons for missing data similar across interventions?	FOR CASE-CONTROLS: Are the proportion of participants and reasons for missing data similar across cases and controls?
Cartsos, 2008[86]	NA-no attrition	NA-no attrition	NA-no attrition	NA-no attrition
Chapurlat, 2013[87]	Overall:0.67 G1: 0 G2: 1.3 Vary by Outcome? No Follow-up Overall: Unclear G1: Unclear G2: Unclear	No	Yes	NA-no attrition
Cryer, 2005[88]	Overall:13.7 [%] G1: 13.8 [%] G2: 13.5[%] G3: [%] No	No	Probably yes	NA-not a case-control
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	Patients missing follow-up xray Overall: 379 / 6459 (5.9%) G1: 198 / 3195 (6.2%) G2: 181 /3183 (5.7)	No	Yes	NA-not a case-control
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	Attrition varies by outcome, lowest for fractures: 475/7868 (6.03%) G1: 231/3933 (5.87%) G2: 244/3935 (6.20%)	No	Yes	NA-not a cohort
Eisman, 2004[96]	Overall: 6.2 [%] G1: 8.0 [%] G2: 4.5 [%] Vary by Outcome? No	No	Probably yes	NA-not a case-control
Fogelman, 2000[97]	G1: 40/179=22% G2: 37/180 =21%	Yes	Yes	NA-not a case control
Greenspan, 2002[98]	Overall: 6.9% G1: 6.3% G2: 7.5% Vary by Outcome? No	No	Yes	NA-not a case-control

© 2018 American Medical Association. All rights reserved.

660 **eTable 54. KQ 4 and 5 risk of bias assessment: Part 6, Missing Data (continued)**

First Author, Year	FOR RCTS and COHORTS: What was the overall attrition? What was the attrition by group? Did attrition vary for different outcomes?	FOR RCTS and COHORTS: Did the study have high attrition raising concern for bias?	FOR RCTS and COHORTS: Are the proportion of participants and reasons for missing data similar across interventions?	FOR CASE-CONTROLS: Are the proportion of participants and reasons for missing data similar across cases and controls?
Greenspan, 2003[99]	Overall: 10[%] G1: 8.6% G2: 9.7 % G3: 9.6 % G4: 10.8% Vary by Outcome? No	No	Yes	NA-not a case-control
Grey, 2010[100]	Overall: 2 [%] G1: 4 [%] G2: 0 [%] Vary by Outcome? No Information	No	Yes	NA-not a case-control
Hosking, 2003[101]	Attrition was only reported at 3 months. G1: 19.8% G2: 21.5% G2: 17.6%	No	No	NA-not a case control
Hosking, 2003[101]	Overall: 25[%] G1: 21.5 [%] G2: 19.8 [%] G3: 17.6 [%] Vary by Outcome? Yes Clincal AE leading to discontinuation: Overall: 17 [%] G1: 14.1 [%] G2: 14.0 [%] G3: 11.1 [%] **of note these are attrition % at 3 months. The study went on for 12 months, b	Yes	Yes	No
Johnell, 2002[102]	Overall: 17%; differences by group NR	No	Yes	NA-no attrition
Keech, 2005[103]	Overall: NR G1: 29% G2: 33% Vary by Outcome? No	Yes	No	NA

© 2018 American Medical Association. All rights reserved.

661 **eTable 54. KQ 4 and 5 risk of bias assessment: Part 6, Missing Data (continued)**

First Author, Year	FOR RCTS and COHORTS: What was the overall attrition? What was the attrition by group? Did attrition vary for different outcomes?	FOR RCTS and COHORTS: Did the study have high attrition raising concern for bias?	FOR RCTS and COHORTS: Are the proportion of participants and reasons for missing data similar across interventions?	FOR CASE-CONTROLS: Are the proportion of participants and reasons for missing data similar across cases and controls?
Kung, 2000[104]	Overall:80 [%] G1: 80[%] G2: 80 [%] G3: [%] Vary by Outcome? No	Yes	Yes	NA-not a case-control
Lasco, 2011[105]	Overall:0	NA-no attrition	NA-no attrition	NA
Lewiecki, 2007[106]	Overall attrition: 5/365=1.00% G1: 0/46 (0%) G2: 5/319 (1.57%)	No	Yes	NA-not a cohort
McCloskey, 2012[107]	Overall:82% G1: NR G2: NR Vary by Outcome? Probably No	No	No information	NA-not a case-control
McClung, 2004[108]	Overall: 16% G1: 15% G2: 13% G3: 18% G4: 17% no	No	Yes	NA-not a case-control
McClung, 2006[109]	Overall:36% G1: 37% G2: 30% G3: 29% G4: 31% Vary by Outcome? No	Yes	No information	NA-not a case-control

662
663

© 2018 American Medical Association. All rights reserved.

664 **eTable 54. KQ 4 and 5 risk of bias assessment: Part 6, Missing Data (continued)**

First Author, Year	FOR RCTS and COHORTS: What was the overall attrition? What was the attrition by group? Did attrition vary for different outcomes?	FOR RCTS and COHORTS: Did the study have high attrition raising concern for bias?	FOR RCTS and COHORTS: Are the proportion of participants and reasons for missing data similar across interventions?	FOR CASE-CONTROLS: Are the proportion of participants and reasons for missing data similar across cases and controls?
McClung, 2006[110]	Overall:10 [%] Not reported by group overal. For below only reported by drug (not dosing group) Vary by Outcome? Yes Withdrawal of consent G1-G7: 8 [%] G8: 2 [%] G9: 7 [%] Adverse effects G1-G7: 2 [%] G8: 0 [%] G9: 2 [%]	No	Yes	NA-not a case-control
McClung, 2009[111]	Overall:90% (Calculated) G1: 91.4% G2: 85.1% G3: 93.1% Vary by Outcome? No	No	no	NA-not a case-control
Meunier, 1999[112]	Overall: 20 of 129 at 24 months (19%), of these, 14 in year 1; differences by group NR	No	Yes	NA-no attrition
Miller, 2008[113]	Overall:[%] 29.7 percent (N=470) discontinued treatment, another 2.9 % (46) failed to return. Additionally the flow chart shows patients who did not complete because of "subject request" and "other."	Yes	Yes	NA-not a case-control
Morii, 2003[114]	Overall: 13%; differences by group NR	No	Yes	NA-no attrition

665
666

© 2018 American Medical Association. All rights reserved.

667 **eTable 54. KQ 4 and 5 risk of bias assessment: Part 6, Missing Data (continued)**

First Author, Year	FOR RCTS and COHORTS: What was the overall attrition? What was the attrition by group? Did attrition vary for different outcomes?	FOR RCTS and COHORTS: Did the study have high attrition raising concern for bias?	FOR RCTS and COHORTS: Are the proportion of participants and reasons for missing data similar across interventions?	FOR CASE-CONTROLS: Are the proportion of participants and reasons for missing data similar across cases and controls?
Murphy, 2001[115]	Overall:15% at 6 mo, 30% at 12 months, 41% at 18 months No data by group G1: [%] G2: [%] G3: [%] Vary by Outcome? No Information	No for 6 months, Yes for 12 and 18 months.	Probably yes	NA-not a case-control
Nakamura, 2012[116]	Overall:8.0 G1: (5/53)9.4 G2: (4/54) 7.4 G3: (5/50)10 G4: (3/55) 5.5 Vary by Outcome? Probably No	No	No information	NA
Orwoll, 2003[117]	Overall:77[17.6%] G1:17 [12%] G2:28 [19%] G3:36 [26%] No Information by outcome	Yes	No	NA-no attrition
Pazianas, 2008[118]	NA-no attrition	NA-no attrition	NA-no attrition	NA-no attrition
Ravn, 1996[119]	Overall: 39/180, 22% G1: 4/30,13% G2: 8/30, 27% G3: 4/30, 13% G4: 6/30, 20% G5: 12/30, 40% G6: 5/30, 17% No	Yes	Yes	NA-not a case-control
Reginster, 2005[120]	Overall: 3% G1: 0 G2: 0 G3: 0 G4: 3% G5: 8% No	No	Yes	NA-not a case-control

668

© 2018 American Medical Association. All rights reserved.

669 **eTable 54. KQ 4 and 5 risk of bias assessment: Part 6, Missing Data (continued)**

First Author, Year	FOR RCTS and COHORTS: What was the overall attrition? What was the attrition by group? Did attrition vary for different outcomes?	FOR RCTS and COHORTS: Did the study have high attrition raising concern for bias?	FOR RCTS and COHORTS: Are the proportion of participants and reasons for missing data similar across interventions?	FOR CASE-CONTROLS: Are the proportion of participants and reasons for missing data similar across cases and controls?
Rhee, 2012[121]	No attrition because data are from registry	NA-no attrition	NA-no attrition	NA-no attrition
Riis, 2001[122]	Overall: 14% G1: 15% G2: 15% G3: 11% no	No	Yes	NA-not a case-control
Samelson, 2014[123]	Overall:82% for the main FREEDOM trial, but this analysis was a subgroup analysis of patients at increased CV risk and with adequate imaging studies. Only 1045 of 2363 patients eligible had evaluation data at baseline and followup. G1: NR G2: NR Vary by	Yes	No information	NA-not a case-control
Shiraki, 2003[124]	G1: 9/56=16% G2: 9/54= 17%	No	No information	NA-not a case control
Simon, 2013[125]	Overall:82% (This is for the overall FREEDOM study; 83% in DXA substudy, 86% in QCT substudy, attrition by treatment group NR) Vary by Outcome? Probably No	No	No information	NA-not a case-control
Sontag, 2010[126]	This article reports only ITT results, but based on original trial, Overall:26% G1: 26% G2: 25% G3: 26% Vary by Outcome? no	Yes	No	NA

670

© 2018 American Medical Association. All rights reserved.

671 **eTable 54. KQ 4 and 5 risk of bias assessment: Part 6, Missing Data (continued)**

First Author, Year	FOR RCTS and COHORTS: What was the overall attrition? What was the attrition by group? Did attrition vary for different outcomes?	FOR RCTS and COHORTS: Did the study have high attrition raising concern for bias?	FOR RCTS and COHORTS: Are the proportion of participants and reasons for missing data similar across interventions?	FOR CASE-CONTROLS: Are the proportion of participants and reasons for missing data similar across cases and controls?
Sorensen, 2008[127]	NA-not an RCT	NA-not an RCT	NA-not an RCT	Yes
Tanko, 2003[128]	Overall: 14% G1: NR G2: NR G3: NR G4: NR G5: NR No	No	Yes	NA-not a case-control
Thiebaud, 1997[129]	Overall: 10% G1: 12.5% (3/24) G2: 3.7% (1/27) G3: 11.5% (3/26) G4: 8.7% (2/23) G5: 7.7% (2/26) Vary by Outcome? No	No	Yes	NA-not a case-control
Tucci, 1996[130]	Overall:29/478=6.0% (from Ns in Table IV) G1: 9.2% G2: 6.4% G3: 8.5% G4: 3.1%	No	No information	NA-not a case-control
Van Staa, 1997[131]	No attrition	NA-no attrition	NA	NA
Vestergaard, 2010[132]	None	No	NA, no attrition	NA, no attrition
Vestergaard, 2011[133]	None	No	NA, no attrition	NA, no attrition
Vestergaard, 2012[134]	None	No	NA, no attrition	NA, no attrition
Vestergaard, 2011[135]	NA, no attrition	NA, no attrition	No	NA-not an RCT
Vestergaard, 2012[136]	None	No	NA, no attrition	NA, no attrition

672 **Abbreviations:** AE=adverse devent; DXA=dual energy x-ray absorptiometry; FREEDOM=Fracture Reduction Evaluation of Denosumab in Osteoporosis Every 6 Month;
673 G=group; KQ=key question; NA=not applicable; NR=not reported; QCT=Quantitative computed tomography; RCT=randomized controlled trials.

674

© 2018 American Medical Association. All rights reserved.

675 **eTable 55. KQ 4 and 5 risk of bias assessment: Part 7, Missing Data**

First Author, Year	FOR ALL STUDIES: Were Appropriate Statistical Methods Used to Account for Missing Data?	Bias Arising From Missing Outcome Data?	Comments
Abrahamsen, 2010[79]	Yes	No	NR
Adachi, 2009[80]	No information	Probably no	Authors do not specifically say they perform an intention to treat analysis.
Barrett-Connor, 2002[81]	Yes	Probably yes	NR
Barrett-Connor, 2004[82]	Yes	No	NR
Bone, 2000[83]	Yes	Probably yes	There was >20% attrition, and over 30% attrition in one of the arms.
Bone, 2008[84]	Yes	No	NR
Boonen, 2012[85]	Yes	No	NR
Cartsos, 2008 [86]	NA-no attrition	No information	No mention of how missing data was handled
Chapurlat, 2013[87]	Yes	Probably no	NR
Cryer, 2005[88]	Yes	Probably no	There is a small difference in reasons for discontinuation. More patients in placebo dropped out due to any clinical AE, however this difference is judged to be small.
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	Yes	No	Missing data=missing xray at follow-up FIT1 (Black, 1996) Overall: 81 / 2027=4.0% G1: 41 / 981=4.2% G2: 40 / 965=4.1% FIT2 (Cummings, 1998 (8400)) Overall: 298/4432 (6.7%) G1: 157 / 2214 (7.1%) G2: 141 / 2218 (6.4%) Combining FIT1 and FIT2
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	Yes	No	NR
Eisman, 2004[96]	Yes	Probably no	More withdrawals for clinical AE in alendronate group vs placebo, but no testing. Results show no difference in discontinuation for UGI AEs
Fogelman, 2000[97]	Yes	Probably no	NR

676

© 2018 American Medical Association. All rights reserved.

677 **eTable 55. KQ 4 and 5 risk of bias assessment: Part 7, Missing Data (continued)**

First Author, Year	FOR ALL STUDIES: Were Appropriate Statistical Methods Used to Account for Missing Data?	Bias Arising From Missing Outcome Data?	Comments
Greenspan, 2002[98]	Yes	No	NR
Greenspan, 2003[99]	Yes	No	ITT analysis
Grey, 2010[100]	Yes	No	NR
Hosking, 2003[101]	Unclear	Probably yes	Unclear what attrition was at 12 months.
Hosking, 2003[101]	Yes	No information	NR
Johnell, 2002[102]	Yes	Probably no	NR
Keech, 2005[103]	Yes	Probably yes	NR
Kung, 2000[104]	Yes	Probably yes	NR
Lasco, 2011[105]	NA-no attrition	Probably no	NR
Lewiecki, 2007[106]	No information	no	NR
McCloskey, 2012[107]	Probably yes	Probably no	NR
McClung, 2004[108]	Yes	No	NR
McClung, 2006[109]	Yes	Probably yes	17, not a case control; overall attrition a little high
McClung, 2006[110]	Yes	No	NR
McClung, 2009[111]	Probably no	unclear	Risk of bias for harms data because it is limited to ITT analysis.
Meunier, 1999[112]	Yes	No	NR
Miller, 2008[113]	Yes	probably no	NR
Morii, 2003[114]	Yes	Probably no	NR
Murphy, 2001[115]	Probably yes	Probably yes	Per protocol analysis probably okay for harms outcomes. Table 6 suggests similar AE profile, but reasons for discontinuation not provided by group.
Nakamura, 2012[116]	Yes	Probably no	NR
Orwoll, 2003[117]	Yes	Probably yes	Differential attrition between arms
Pazianas, 2008[118]	NA-no attrition	No information	No mention of how missing data was handled
Ravn, 1996[119]	No information	Probably no	high overall and differential attrition; however, safety appears to have been collected and reported on a larger subset of the population
Reginster, 2005[120]	Yes	No	NR
Rhee, 2012[121]	NA-no attrition	no	NR
Riis, 2001[122]	Yes	No	NR
Samelson, 2014[123]	No	Probably yes	NR
Shiraki, 2003[124]	Yes	Probably no	NR
Simon, 2013[125]	Yes	Probably no	NR
Sontag, 2010[126]	Yes	Probably yes	NR
Sorensen, 2008[127]	Yes	No	*Authors report Danish registry information is complete.

© 2018 American Medical Association. All rights reserved.

678 **eTable 55. KQ 4 and 5 risk of bias assessment: Part 7, Missing Data (continued)**

First Author, Year	FOR ALL STUDIES: Were Appropriate Statistical Methods Used to Account for Missing Data?	Bias Arising From Missing Outcome Data?	Comments
Tanko, 2003[128]	Yes	Probably no	Unable to calculate group attrition
Thiebaud, 1997[129]	Yes	Probably no	Used ITT but one patient who dropped out before treatment because of inability to administer the drug was not included. Missing values were not replaced.
Tucci, 1996[130]	Yes	Probably no	Study was extended for a third year, 14 participants did not consent to blinded treatment for a third year, 5 declined third year altogether.
Van Staa, 1997[131]	NA	No information	The study did not provide any information on attrition or missing data.
Vestergaard, 2010[132]	NA, no attrition	No	NR
Vestergaard, 2011[133]	NA, no attrition	No	NR
Vestergaard, 2012[134]	NA, no attrition	No	NR
Vestergaard, 2011[135]	NA-not an RCT	No information	NR
Vestergaard, 2012[136]	NA, no attrition	No	NR

679 **Abbreviations:** AE=adverse event; FIT=fracture intervention trial; ITT=intent to treat; KQ=key question; NA=not applicable; NR=not reported; UGI=upper gastrointestinal.

680
681

© 2018 American Medical Association. All rights reserved.

682 **eTable 56. KQ 4 and 5 risk of bias assessment: Part 8, Departure from Intended Intervention**

First Author, Year	FOR RCTs of treatment (NA for screening): Were the patients unaware of their intervention status of participants?	FOR all RCTs: Were the trial personnel and clinicians unaware of the intervention status of participants?	FOR ALL STUDIES Was intervention fidelity adequate?	FOR ALL STUDIES: Did the study have enough cross-overs or contamination that would raise concern for bias?	Bias arising from departures from intended interventions?	Departures from interventions comments
Abrahamsen, 2010[79]	NA-not an RCT	NA-not an RCT	Probably yes	No information	Probably no	NR
Adachi, 2009[80]	Yes	Yes	No information	No information	Probably no	No data on adherence
Barrett-Connor, 2002[81]	Yes	Yes	Probably yes	Probably no	Probably no	In year 4 could take additional medications.
Barrett-Connor, 2004[82]	Yes	Yes	Yes	No	No	stated in larger study that 92% of women took more than 80% of study medication
Bone, 2000[83]	Yes	Yes	No information	No information	No	The authors did not report crossover, but were thorough about patient accounting
Bone, 2008[84]	Probably no	Probably no	NA (subcutaneous)	No information	Probably no	NR
Boonen, 2012[85]	Probably no	Probably no	NA (subcutaneous)	No information	Probably no	NR
Cartsos, 2008 [86]	NA-not an RCT	NA-not an RCT	No information	No information	No information	Fidelity, not sure if participants took medication correctly; no information on cross-overs but not clear if other treatments were allowed
Chapurlat, 2013[87]	Yes	Yes	Yes	No	No	NR
Cryer, 2005[88]	Yes	Yes	Yes	No	No	
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	Yes	Yes	Yes	No	No	NR
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	Probably no	Probably no	NA (subcutaneous)	No information	Probably no	NR

© 2018 American Medical Association. All rights reserved.

683 **eTable 56. KQ 4 and 5 risk of bias assessment: Part 8, Departure from Intended Intervention (continued)**

First Author, Year	FOR RCTs of treatment (NA for screening): Were the patients unaware of their intervention status of participants?	FOR all RCTs: Were the trial personnel and clinicians unaware of the intervention status of participants?	FOR ALL STUDIES Was intervention fidelity adequate?	FOR ALL STUDIES: Did the study have enough cross-overs or contamination that would raise concern for bias?	Bias arising from departures from intended interventions?	Departures from interventions comments
Eisman, 2004[96]	NR	Yes	Yes	No	No	Mean compliance 95% and 96% for alendronate and placebo groups
Fogelman, 2000[97]	Yes	Yes	No information	Probably no	Probably no	NR
Greenspan, 2002[98]	Yes	Yes	Probably yes	Probably no	Probably no	NR
Greenspan, 2003[99]	Yes	Yes	Yes	No	No	NR
Grey, 2010[100]	Yes	Yes	Yes	No	No	NR
Hosking, 2003[101]	Yes	Yes	Yes	No	No	NR
Hosking, 2003[101]	Yes	Yes	Yes	No information	No	>75% adherence to medications
Johnell, 2002[102]	Yes	Yes	Yes	No	Probably no	NR
Keech, 2005[103]	Yes	Yes	Yes	No	No	NR
Kung, 2000[104]	Yes	Yes	No information	No information	Probably no	NR
Lasco, 2011[105]	NA-not an RCT	NA-not an RCT	Probably yes	Probably no	Probably no	NR
Lewiecki, 2007[106]	Probably no	Probably no	NA (subcutaneous)	No information	Probably no	NR
McCloskey, 2012[107]	Yes	Yes	Yes	No information	No	NR
McClung, 2004[108]	Yes	Yes	Yes	No	No	Compliance in mid to high 80s
McClung, 2006[109]	Yes	Yes	No information	No information	probably yes	Adherence unknown
McClung, 2006[110]	Yes	Yes	Yes	No	No	Of note - double blinding for denosumab but NOT alendronate (open label) all answers are for denosumab. For alendronate (no, no, yes, no information, probably yes)

© 2018 American Medical Association. All rights reserved.

684 **eTable 56. KQ 4 and 5 risk of bias assessment: Part 8, Departure from Intended Intervention (continued)**

First Author, Year	FOR RCTs of treatment (NA for screening): Were the patients unaware of their intervention status of participants?	FOR all RCTs: Were the trial personnel and clinicians unaware of the intervention status of participants?	FOR ALL STUDIES Was intervention fidelity adequate?	FOR ALL STUDIES: Did the study have enough cross-overs or contamination that would raise concern for bias?	Bias arising from departures from intended interventions?	Departures from interventions comments
McClung, 2009[111]	Yes	Yes	Probably yes	Probably no	Probably no	NR
Meunier, 1999[112]	Yes	Yes	Probably yes	No	No	NR
Miller, 2008[113]	Yes	probably yes	No information	No information	Probably no	NR
Morii, 2003[114]	Yes	No information	Probably yes	No	No	NR
Murphy, 2001[115]	Yes	Yes	Yes	No	No	Only 4 patients failed to take >75% of assigned drug
Nakamura, 2012[116]	Probably yes	Probably yes	Yes	No	No	NR
Orwoll, 2003[117]	Yes	Yes	Yes	Probably no	Probably no	Patient-administered injections of placebo or drug
Pazianas, 2008[118]	NA-not an RCT	NA-not an RCT	No information	No information	No information	Fidelity, not sure if participants took medication correctly; no information on cross-overs but not clear if other treatments were allowed
Ravn, 1996[119]	Yes	No	No information	No	Probably no	Data safety review committee (DSRC) was not blinded to treatment, and they monitored adverse events during each step. Information on compliance was not provided
Reginster, 2005[120]	NA-not an RCT	NA-not an RCT	Probably no	No	Probably yes	No way to determine if participants took dose
Rhee, 2012[121]	Yes	Yes	Yes	No	No	NR
Riis, 2001[122]	Probably yes	Probably yes	Yes	No	Probably no	NR

© 2018 American Medical Association. All rights reserved.

685 **eTable 56. KQ 4 and 5 risk of bias assessment: Part 8, Departure from Intended Intervention (continued)**

First Author, Year	FOR RCTs of treatment (NA for screening): Were the patients unaware of their intervention status of participants?	FOR all RCTs: Were the trial personnel and clinicians unaware of the intervention status of participants?	FOR ALL STUDIES Was intervention fidelity adequate?	FOR ALL STUDIES: Did the study have enough cross-overs or contamination that would raise concern for bias?	Bias arising from departures from intended interventions?	Departures from interventions comments
Samelson, 2014[123]	Yes	Yes	Probably yes	No information	Probably no	NR
Shiraki, 2003[124]	Yes	Yes	No information	Probably no	Probably no	NR
Simon, 2013[125]	Yes	Yes	Probably yes	No information	No	NR
Sontag, 2010[126]	Probably yes	Probably yes	No	Probably no	Probably no	The study was reported as double-blind but no other details were provided. The placebo arm received active treatment after 1 year but the results are not reported separately for before and after receipt of active treatment
Sorensen, 2008[127]	NA-not an RCT	NA-not an RCT	Probably yes	No information	Probably no	NR
Tanko, 2003[128]	Yes	Yes	No information	No	Probably no	A large proportion of patients in each study group took at least 75% of study medication: 89% placebo, 88.8% (5 mg), 90.1% (10 mg) and 88.7% (20 mg) patients.
Thiebaud, 1997[129]	Yes	No	No information	No	Probably no	Information on compliance was not provided Investigator was not blind for all arms
Tucci, 1996[130]	Yes	Yes	Yes	No	No	Investigators only evaluated blinded results (excluded patients who declined blinding for third year)

© 2018 American Medical Association. All rights reserved.

686 **eTable 56. KQ 4 and 5 risk of bias assessment: Part 8, Departure from Intended Intervention (continued)**

First Author, Year	FOR RCTs of treatment (NA for screening): Were the patients unaware of their intervention status of participants?	FOR all RCTs: Were the trial personnel and clinicians unaware of the intervention status of participants?	FOR ALL STUDIES Was intervention fidelity adequate?	FOR ALL STUDIES: Did the study have enough cross-overs or contamination that would raise concern for bias?	Bias arising from departures from intended interventions?	Departures from interventions comments
Van Staa, 1997[131]	NA-not an RCT	NA-not an RCT	No information	No	No information	Did not evaluate adherence
Vestergaard, 2010[132]	NA-not an RCT	NA-not an RCT	No information	No information	No information	NR
Vestergaard, 2011[133]	NA-not an RCT	NA-not an RCT	No information	No information	No information	NR
Vestergaard, 2012[134]	NA-not an RCT	NA-not an RCT	No information	No information	No information	NR
Vestergaard, 2011[135]	No information	NA-no benefits outcomes	NA-no benefits outcomes	NA-no benefits outcomes	Probably no	NR
Vestergaard, 2012[136]	NA-not an RCT	NA-not an RCT	No information	No information	No information	NR

687 **Abbreviations:** KQ=key question; NA=not applicable;

688
689

© 2018 American Medical Association. All rights reserved.

690 **eTable 57. KQ 4 and 5 risk of bias assessment: Part 9, Outcome Measurement**

First Author, Year	FOR ALL STUDIES: Were benefit outcomes (e.g., fractures) adequately described, pre-specified, valid, and reliable?	FOR ALL STUDIES: Were Similar Techniques Used Among Groups to Ascertain Harm Outcomes?	FOR ALL STUDIES: Was the Duration of Follow-Up Adequate to Assess Harm Outcomes?	Bias Arising from Measurement of Outcomes?	Comments
Abrahamsen, 2010[79]	NA-no benefits outcomes	Probably yes	Probably yes	Probably yes	Not able to identify atypia.
Adachi, 2009[80]	NA-no benefits outcomes	Yes	Yes	Probably no	There was not specific information about how often patient's assessed for harms, though did describe adequate blinding of patients.
Barrett-Connor, 2002[81]	NA-no benefits outcomes	Yes	Yes	No	NR
Barrett-Connor, 2004[82]	Yes	Yes	Yes	No	NR
Bone, 2000[83]	Probably Yes	Probably yes	Yes	Probably yes	Report that patients were seen at 3, 6, 12, 18, 24 months, but don't specifically describe clinical assessment (i.e. patient assessed for harms at this time)
Bone, 2008[84]	Yes	Yes	Probably yes	Probably no	NR
Boonen, 2012[85]	Yes	Yes	Yes	No	NR
Cartsos, 2008 [86]	NA-no benefits outcomes	Probably yes	Probably yes	Probably yes	Not clear how outcomes were measured due to only a code being provided
Chapurlat, 2013[87]	Yes	Probably yes	Yes	No	NR
Cryer, 2005[88]	Yes	Yes	Yes	No	NR
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	Yes	Yes	Yes	No	NR
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	Yes	Yes	Probably yes	Probably no	NR

691

© 2018 American Medical Association. All rights reserved.

692 **eTable 57. KQ 4 and 5 risk of bias assessment: Part 9, Outcome Measurement (continued)**

First Author, Year	FOR ALL STUDIES: Were benefit outcomes (e.g., fractures) adequately described, pre-specified, valid, and reliable?	FOR ALL STUDIES: Were Similar Techniques Used Among Groups to Ascertain Harm Outcomes?	FOR ALL STUDIES: Was the Duration of Follow-Up Adequate to Assess Harm Outcomes?	Bias Arising from Measurement of Outcomes?	Comments
Eisman, 2004[96]	Yes	Yes	Yes	No	NR
Fogelman, 2000[97]	Probably Yes	Yes	Yes	Probably no	NR
Greenspan, 2002[98]	NA-no benefits outcomes	Yes	Yes	No	NR
Greenspan, 2003[99]	Yes	Yes	Yes	No	NR
Grey, 2010[100]	Probably yes	Probably yes	Yes	Probably no	Looked at parent article to identify clinical assessment of harms - no information. The antiresorptive effects of a single dose of zoledronate persist for two years: a randomized, placebo-controlled trial in osteopenic postmenopausal women.
Hosking, 2003[101]	NA-no benefits outcomes	Yes	Probably yes	Probably no	NR
Johnell, 2002[102]	NA-no benefits outcomes	Yes	Probably yes	Probably no	12 month study
Keech, 2005[103]	NA-no benefits outcomes	Yes	Yes	No	NR
Kung, 2000[104]	NA-no benefits outcomes	Yes	Yes	Probably yes	No information on how harms ascertained
Lasco, 2011[105]	NA-no benefits outcomes	No information	No information	Probably no	NR
Lewiecki, 2007[106]	Yes	Yes	Probably yes	Probably no	NR
McCloskey, 2012[107]	Yes	NA-no harms outcomes	NA-no harms outcomes	No	NR
McClung, 2004[108]	NA-no benefits outcomes	Yes	Yes	No	NR
McClung, 2006[109]	NA-no benefits outcomes	Yes	Yes	No	NR
McClung, 2006[110]	Yes	Yes	Yes	No	NR
McClung, 2009[111]	NA-no benefits outcomes	yes	yes	Probably no	NR
Meunier, 1999[112]	NA-no benefits outcomes	Yes	Yes	Probably no	Followup was 2 years
Miller, 2008[113]	NA-no benefits outcomes	Yes	Yes	Probably no	NR
Morii, 2003[114]	NA-no benefits outcomes	Yes	Yes	Probably no	NR
Murphy, 2001[115]	Yes	Yes	Yes	No	NR
Nakamura, 2012[116]	Yes	Yes	Probably yes	No	NR

© 2018 American Medical Association. All rights reserved.

693 **eTable 57. KQ 4 and 5 risk of bias assessment: Part 9, Outcome Measurement (continued)**

First Author, Year	FOR ALL STUDIES: Were benefit outcomes (e.g., fractures) adequately described, pre-specified, valid, and reliable?	FOR ALL STUDIES: Were Similar Techniques Used Among Groups to Ascertain Harm Outcomes?	FOR ALL STUDIES: Was the Duration of Follow-Up Adequate to Assess Harm Outcomes?	Bias Arising from Measurement of Outcomes?	Comments
Orwoll, 2003[117]	Yes	Yes	Probably no	Probably no	NR
Pazianas, 2008[118]	NA-no benefits outcomes	Yes	Probably yes	Probably no	NR
Ravn, 1996[119]	NA-no benefits outcomes	Yes	Yes	No	NR
Reginster, 2005[120]	NA-no benefits outcomes	Yes	Yes	No	NR
Rhee, 2012[121]	NA-no benefits outcomes	Yes	Yes	No	NR
Riis, 2001[122]	NA-no benefits outcomes	Yes	Yes	No	NR
Samelson, 2014[123]	NA-no benefits outcomes	Yes	Yes	Probably yes	Post hoc analysis and the approach to reporting cardiovascular events in this analysis is different from reporting in the main FREEDOM trial where cardiovascular events were adjudicated by a panel.
Shiraki, 2003[124]	NA-no benefits outcomes	Yes	Yes	Probably yes	NR
Simon, 2013[125]	Probably Yes	NA-no harms outcomes	NA-no harms outcomes	Probably no	NR
Sontag, 2010[126]	Yes	Yes	Yes	Probably no	NR
Sorensen, 2008[127]	NA-no benefits outcomes	Yes	Probably yes	Probably no	Case control - harms only identified in the case group
Tanko, 2003[128]	NA-no benefits outcomes	Yes	Yes	No	NR
Thiebaud, 1997[129]	NA-no benefits outcomes	Yes	Yes	No	NR
Tucci, 1996[130]	Yes	Yes	Yes	No	Of note, there are some data on reduction of vertebral fractures, but investigators have planned another arm with future reporting. This study not powered for fracture reduction.
Van Staa, 1997[131]	NA-no benefits outcomes	Yes	Yes	No	NR
Vestergaard, 2010[132]	NA-no benefits outcomes	Yes	Probably yes	Probably no	NR

© 2018 American Medical Association. All rights reserved.

694 **eTable 57. KQ 4 and 5 risk of bias assessment: Part 9, Outcome Measurement (continued)**

First Author, Year	FOR ALL STUDIES: Were benefit outcomes (e.g., fractures) adequately described, pre-specified, valid, and reliable?	FOR ALL STUDIES: Were Similar Techniques Used Among Groups to Ascertain Harm Outcomes?	FOR ALL STUDIES: Was the Duration of Follow-Up Adequate to Assess Harm Outcomes?	Bias Arising from Measurement of Outcomes?	Comments
Vestergaard, 2011[133]	NA-no benefits outcomes	Yes	Probably yes	Probably no	NR
Vestergaard, 2012[134]	NA-no benefits outcomes	yes	Probably yes	Probably yes	NR
Vestergaard, 2011[135]	Probably Yes	Probably no	Yes	Probably yes	NR
Vestergaard, 2012[136]	NA-no benefits outcomes	yes	Probably yes	Probably yes	NR

695 **Abbreviations:** FREEDOM=Fracture Reduction Evaluation of Denosumab in Osteoporosis Every 6 Month; KQ=key question; NA=not applicable; NR=not reported;

696
697

© 2018 American Medical Association. All rights reserved.

698 **eTable 58. KQ 4 and 5 risk of bias assessment: Part 10, Reporting**

First Author, Year	FOR RCTS AND COHORTS: Is the Reported Effect Estimate Unlikely to be Selected, on the Basis of the Results, From Multiple Outcomes Measurements Within the Domain, Multiple Analyses, or Different Subgroups?	FOR CASE-CONTROL STUDIES: Is the Reported Effect Estimate Unlikely to be Selected, on the Basis of the Results, from Multiple Definitions of the Intervention?	Bias arising From Selection of Reported Results?	Comments
Abrahamsen, 2010[79]	Probably yes	NA-not a case-control	Probably no	NR
Adachi, 2009[80]	Yes	NA-not a case-control	No	NR
Barrett-Connor, 2002[81]	No	NA-not a case-control	No	NR
Barrett-Connor, 2004[82]	No	NA-not a case-control	No	NR
Bone, 2000[83]	Yes	NA-not a case-control	No	NR
Bone, 2008[84]	Probably no	NA-not a case-control	Probably no	NR
Boonen, 2012[85]	Yes	NA-not a case-control	Probably no	NR
Cartsos, 2008 [86]	NA-not an RCT	Probably yes	Probably no	None
Chapurlat, 2013[87]	No	NA-not a case-control	No	NR
Cryer, 2005[88]	Yes	NA-not a case-control	No	NR
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	Yes	NA-not a case-control	No	NR
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	Probably no	NA-not a case-control	Probably no	NR
Eisman, 2004[96]	Yes	NA-not a case-control	No	NR
Fogelman, 2000[97]	Yes	NA-not a case-control	No	NR
Greenspan, 2002[98]	Yes	NA-not a case-control	No	NR
Greenspan, 2003[99]	Yes	NA-not a case-control	No	NR
Grey, 2010[100]	Yes	NA-not a case-control	No	NR
Hosking, 2003[101]	Yes	NA-not a case-control	No	NR
Hosking, 2003[101]	Yes	NA-not a case-control	No	NR
Johnell, 2002[102]	Probably yes	NA-not a case-control	Probably no	NR
Keech, 2005[103]	No	NA-not a case-control	No	NR
Kung, 2000[104]	Yes	NA-not a case-control	No	NR
Lasco, 2011[105]	Probably no	NA-not a case-control	Probably no	NR
Lewiecki, 2007[106]	Probably no	NA-not a case-control	Probably no	NR
McCloskey, 2012[107]	No	NA-not a case-control	Probably no	NR

© 2018 American Medical Association. All rights reserved.

699 **eTable 58. KQ 4 and 5 risk of bias assessment: Part 10, Reporting (continued)**

First Author, Year	FOR RCTS AND COHORTS: Is the Reported Effect Estimate Unlikely to be Selected, on the Basis of the Results, From Multiple Outcomes Measurements Within the Domain, Multiple Analyses, or Different Subgroups?	FOR CASE-CONTROL STUDIES: Is the Reported Effect Estimate Unlikely to be Selected, on the Basis of the Results, from Multiple Definitions of the Intervention?	Bias arising From Selection of Reported Results?	Comments
McClung, 2004[108]	No	No	No	NR
McClung, 2006[109]	No	NA-not a case-control	No	NR
McClung, 2006[110]	Yes	NA-not a case-control	No	Study was powered for primary outcome of urinary markers, not harms. Report nominal p values for harms
McClung, 2009[111]	probably yes	NA-not a case-control	Probably no	NR
Meunier, 1999[112]	Yes	NA-not a case-control	Probably no	NR
Miller, 2008[113]	Probably no	NA-not a case-control	Probably no	NR
Morii, 2003[114]	Yes	NA-not a case-control	Probably no	NR
Murphy, 2001[115]	Yes	NA-not a case-control	No	NR
Nakamura, 2012[116]	No	NA-not a case-control	No	NR
Orwoll, 2003[117]	Probably yes	NA-not a case-control	Probably no	NR
Pazianas, 2008[118]	NA-not an RCT	Probably yes	Probably no	NR
Ravn, 1996[119]	No	No	No	NR
Reginster, 2005[120]	No	No	No	NR
Rhee, 2012[121]	No	NA-not a case-control	No	NR
Riis, 2001[122]	No	No	No	NR
Samelson, 2014[123]	Probably yes	NA-not a case-control	No	It is not clear how the cardiovascular adverse events reported in this study relate to the harms provided in the main FREEDOM trial. This appears to be a post-hoc analysis.
Shiraki, 2003[124]	Yes	NA-not a case-control	No	NR
Simon, 2013[125]	Probably yes	NA-not a case-control	Probably no	NR
Sontag, 2010[126]	Probably no	NA-not a case-control	Probably no	NR
Sorensen, 2008[127]	NA-not an RCT	Yes	No	NR
Tanko, 2003[128]	No	No	No	NR
Thiebaud, 1997[129]	No	No	No	NR
Tucci, 1996[130]	Yes	NA-not a case-control	No	Stepwise Tukey trend test to adjust for multiple comparisons

© 2018 American Medical Association. All rights reserved.

700 **eTable 58. KQ 4 and 5 risk of bias assessment: Part 10, Reporting (continued)**

First Author, Year	FOR RCTS AND COHORTS: Is the Reported Effect Estimate Unlikely to be Selected, on the Basis of the Results, From Multiple Outcomes Measurements Within the Domain, Multiple Analyses, or Different Subgroups?	FOR CASE-CONTROL STUDIES: Is the Reported Effect Estimate Unlikely to be Selected, on the Basis of the Results, from Multiple Definitions of the Intervention?	Bias arising From Selection of Reported Results?	Comments
Van Staa, 1997[131]	Yes	NA-not a case-control	No	Intervention status defined as patients who had received a prescription; adherence not measured; attrition and how missing data was handled was not reported
Vestergaard, 2010[132]	Probably no	NA-not a case-control	Probably no	NR
Vestergaard, 2011[133]	Probably no	NA-not a case-control	Probably no	NR
Vestergaard, 2012[134]	Probably no	NA-not a case-control	Probably no	NR
Vestergaard, 2011[135]	Probably no	NA-not a case-control	Probably no	NR
Vestergaard, 2012[136]	Probably no	NA-not a case-control	Probably no	NR

701 **Abbreviations:** FREEDOM=Fracture Reduction Evaluation of Denosumab in Osteoporosis Every 6 Month; KQ=key question; NA=not applicable; NR=not reported;
702 RCT=randomized controlled trials.

703

© 2018 American Medical Association. All rights reserved.

704 **eTable 59. KQ 4 and 5 risk of bias assessment: Part 11, Overall Ratings**

First Author, Year	Rating Overall (if you rate one of the domains as having bias, the study cannot be high quality).	Rating Justification	Does Quality Rating of Study Vary by Outcome?
Abrahamsen, 2010[79]	Poor	Risk of bias from residual confounding and measurement of outcomes	No
Adachi, 2009[80]	Fair	Baseline differences between groups raise some concerns for risk of bias.	No
Barrett-Connor, 2002[81]	Fair	About 25% lost to follow-up. Also year four data allows additional therapy for osteoporosis which was different per group though small number (<7%) - this study included year four partcipants but didn't report concomittant medicationss. Additionally, there was differential loss to follow-up due to excessive bone loss in the placebo group (3% vs 1%).	No
Barrett-Connor, 2004[82]	Fair	About 25% lost to follow-up. Also year four data allows additional therapy for osteoporosis which was different per group though small number (<7%) - this study included year four partcipants but didn't report concomittant meds. (No sensitivity analysis looking at three years of data where no additional meds.) Additionally, there was differential loss to follow-up due to excessive bone loss in the placebo group (3% vs 1%).	No
Bone, 2000[83]	Poor	High attrition and no information about how harms were specified or assessed.	No
Bone, 2008[84]	Fair	Some uncertainties in reporting of randomization, allocation concealment, blinding	no
Boonen, 2012[85]	Good	NR	No
Cartsos, 2008[86]	Poor	not clear how outcomes were measured. fidelity, not sure if participants took medication correctly; no information on cross-overs but not clear if other treatments were allowed No mention of how missing data was handled sample not representative of total population intervention based on dispensing information from claims data, information on dose etc. not available	No
Chapurlat, 2013[87]	Fair	Considering IVR with minimization scheme to be just adequate and unclear way drop outs handled.	No
Cryer, 2005[88]	Good	fair for differential attrition, no information on contamination	No
Cummings, 1998[89] Quandt, 2005[90] Bauer, 2000[91]	Good	NR	No

705
706

© 2018 American Medical Association. All rights reserved.

707 **eTable 59. KQ 4 and 5 risk of bias assessment: Part 11, Overall Ratings (continued)**

First Author, Year	Rating Overall (if you rate one of the domains as having bias, the study cannot be high quality).	Rating Justification	Does Quality Rating of Study Vary by Outcome?
Cummings, 2009[92]; Watts, 2012[93]; McClung, 2012[94]; Boonen, 2011[95]	Fair	Some uncertainties in reporting of randomization, allocation concealment, blinding	no
Eisman, 2004[96]	Good	NR	No
Fogelman, 2000[97]	Fair	NR	No
Greenspan, 2002[98]	Fair	missing info on randomization. Also no washout period for patients previously on bisphosphonates.	No
Greenspan, 2003[99]	Good	NR	No
Grey, 2010[100]	Fair	Differences in baseline fracture rates, minimal specification of harm outcomes.	No
Hosking, 2003[101]	Fair	NR	No
Hosking, 2003[101]	Fair	Fair or Poor depending on how rate Attirtion module	No
Johnell, 2002[102]	Good	NR	No
Keech, 2005[103]	Fair	About 25% lost to follow-up. Also year four data allows additional therapy for osteoporosis which was different per group though small number (<7%) - this study included year four partcipants but didn't report concomittant medications. (No sensitivity analysis looking at three years of data where no additional meds.) Additionally, there was differential loss to follow-up due to excessive bone loss in the placebo group (3% vs 1%).	No
Kung, 2000[104]	Poor	No information on randomization methods, fidelity, contamination, 20% attrition with not enough info to judge differential attrition, and poorly specified harms outcomes (very specirfic patient self-reported adverse experiences, with no indication as to seriousness of AE, whether the AE resulted in discontinuation, and further, the data offered is number of events, not number of women, making it difficult to know whether the risk is higher in one group, compared to the other.	No
Lasco, 2011[105]	Poor	Potential for confounding	No
Lewiecki, 2007[106]	Fair	Some uncertainties in reporting of randomization, allocation concealment, blinding	No
McCloskey, 2012[107]	Fair	No detail on randomization and allocation concealment prevents this from being rated as Good. No fatal flaws	No
McClung, 2004[108]	Fair	No information provided on method of randomization or concealment	No
McClung, 2006[109]	Fair	Overall attrition high, not a lot of information provided on randomization process; Fidelity issue: no information on if participants actually took their assigned doses	No

© 2018 American Medical Association. All rights reserved.

708 **eTable 59. KQ 4 and 5 risk of bias assessment: Part 11, Overall Ratings (continued)**

First Author, Year	Rating Overall (if you rate one of the domains as having bias, the study cannot be high quality).	Rating Justification	Does Quality Rating of Study Vary by Outcome?
McClung, 2006[110]	Good	Good for denosumab. For alendronate Poor for lack of blinding.	No
McClung, 2009[111]	Fair	Higher risk of bias for harms than benefits (ITT analysis understates harms)	No
Meunier, 1999[112]	Good	Documentation on randomization missing, outcomes mostly self report	no
Miller, 2008[113]	Fair	Not possible to say how missing cases were accounted for in the analysis. Study has a potential to underestimate harms by using N randomized in the denominator and N retained in the numerator.	No
Morii, 2003[114]	Fair	NR	No
Murphy, 2001[115]	Poor	Very poor attrition at 12 and 18 months, and unable to assess differential attrition, missing information on randomization	No
Nakamura, 2012[116]	Fair	The article was lacking information on method of randomization and concealment; lack of information on those who discontiuned study	No
Orwoll, 2003[117]	fair	Differential attrition; higher in treatment arm; used ITT to adjust for analysis	No
Pazianas, 2008[118]	Poor	fidelity, not sure if participants took medication correctly; no information on cross-overs but not clear if other treatments were allowed No mention of how missing data was handled sample not representative of total population intervention based on dispensing information from claims data, information on dose etc. not available	No
Ravn, 1996[119]	Fair	High attrition, however, safety appears to have been collected and reported on a larger subset of the population. No information provided on method of randomization or concealment.	No
Reginster, 2005[120]	Fair	No information provided on method of randomization or concealment Information on compliance was not provided	No
Rhee, 2012[121]	Poor	Potential bias arising from creation of an new user cohort and from restriction to those without switches	No
Riis, 2001[122]	Fair	No information provided on method of randomization or concealment	No
Samelson, 2014[123]	Poor	No detail on randomization and allocation concealment prevents the main trial from being rated as Good. Attrition/missing data and outcome measurement in this specific sub-study make this analysis high risk of bias, thus Poor Quality.	No

© 2018 American Medical Association. All rights reserved.

709 **eTable 59. KQ 4 and 5 risk of bias assessment: Part 11, Overall Ratings (continued)**

First Author, Year	Rating Overall (if you rate one of the domains as having bias, the study cannot be high quality).	Rating Justification	Does Quality Rating of Study Vary by Outcome?
Shiraki, 2003[124]	Fair	NR	No
Simon, 2013[125]	Fair	In the end, the only outcome that is of interest are wrist fractures in subgroups based on baseline risk.	No
Sontag, 2010[126]	Poor	The open label portion of the trial allowed patient choice, and as result, outcomes could be result of confounding because of prognostic variables	No
Sorensen, 2008[127]	Good	NR	No
Tanko, 2003[128]	Fair	No information provided on method of randomization or concealment not able to calculate group attrition	No
Thiebaud, 1997[129]	Fair	No information provided on method of randomization or concealment Slight differences length of menopause Information on compliance was not provided Investigator was not blind for all arms	No
Tucci, 1996[130]	Fair	Randomization methods, fidelity, contamination missing info.	No
Van Staa, 1997[131]	Poor	NR	No
Vestergaard, 2010[132]	Poor	Concerns include lack of adjustment for all potential confounders, particularly OTC NSAID use and smoking. Additionally, the study does not control for adherence.	no
Vestergaard, 2011[133]	Poor	Concerns include lack of adjustment for all potential confounders. For example, smoking, hypertension, diabetes could explain the stroke, and it's possible that these underlying conditions are highly associated with both the osteoporosis medications and the outcome.	No
Vestergaard, 2012[134]	Poor	Concerns include lack of adjustment for all potential confounders. For example, smoking and hypertension could explain the stroke, and it's possible that these underlying conditions are highly associated with both the osteoporosis medications and the outcome.	No
Vestergaard, 2011[135]	Poor	Concerns include lack of adjustment for all potential confounders, particularly underlying disease that might also be related to the choice of medication for steoporosis and the outcome. Additionally the outcome did not distinguish between typical and typical fractures.	No
Vestergaard, 2012[136]	Poor	Concerns include lack of adjustment for all potential confounders, particularly underlying causes of inflammatory jaw disease (e.g., autoimmune disorders) that might also be related to risk factors for osteoporosis. Additionally the outcome includes many varied conditions with different etiologies that might be unrelated to osteoporosis.	No

710 **Abbreviations:** AE=adverse event; ITT=intent to treat; IVR=interactive voice response; KQ=key question; NR=not reported; NSAIDS=nonsteroidal anti-inflammatory drugs;
711 OTC=over the counter.

© 2018 American Medical Association. All rights reserved.

712 **eTable 60. Characteristics and accuracy of clinical risk assessment tools in identifying osteoporosis (KQ 2a)**

Instrument (Instrument components)	Clinical and Geographic Setting	Mean Age (Years)	Sex	Race/ Ethnicity	Area under the curve			Sensitivity and specificity				
					No. of Studies	No. of Partici-pants	Pooled AUC (95% CI) or Range of AUCs;[a]	No. of Studies	No. of Partici-pants	Range of Sensitivity	Range of Specificity	Threshold[b]
ABONE[12, 17] (Age, body size, no estrogen use for at least 6 months)	General population; Canada Singapore	66. to 68.4	All women	White and Chinese	2	2500	0.70 to 0.72	1	2365	83.3 (78.5-88.0)	47.7 (45.6-49.8)	≥2
AMMEB[20, 21] (Age, BMI, age at menarche, postmenopausal period)	General practices; Italy	65	All women	NR	2	1520	0.63 to 0.71	NR	NR	NR	NR	NR
DOEScore[37] (Age, body weight, and history of fracture)	Population-based cohort; Dubbo, Australia	70.5	All women	98.6% Caucasian; 1.4% Aboriginal (overall cohort, NR for included sample)	1	410	0.75 (0.691 to 0.809)	1	410	82% (NR)	52% (NR)	>10
FRAX without BMD for 10-year risk of hip fracure[27, 40] (Age, race, rheumatoid arthritis, history of prior fracture, medication use, smoking, alcohol intake, and parental history of hip fracture)	General practice, Spain[27]	61	All women	100% Caucasian	1	505	0.82 (NR)	1	505	NR	NR	NR
	General practice; Australia[40]	78.2	45.1% women	Not reported	1	626	0.70 (0.64 to 0.75)	1	626	92.2 (NR)	37.7 (NR)	Hip≥3%

713
714

© 2018 American Medical Association. All rights reserved.

eTable 60. Characteristics and accuracy of clinical risk assessment tools in identifying osteoporosis (KQ 2a) (continued)

					Area under the curve			Sensitivity and specificity				
Instrument (Instrument components)	**Clinical and Geographic Setting**	**Mean Age (Years)**	**Sex**	**Race/ Ethnicity**	**No. of Studies**	**No. of Partici-pants**	**Pooled AUC (95% CI) or Range of AUCs;[a]**	**No. of Studies**	**No. of Partici-pants**	**Range of Sensitivity**	**Range of Specificity**	**Threshold[b]**
FRAX without BMD for 10-year risk of major osteoporotic fracure[9, 16, 19, 27, 30, 40]	General practice; USA, Spain; Population-based cohort, Canada[9, 19, 27, 30]	57 to 57.7	All women	72% white, 17% black, 8% Hispanic in one study[19]; 97%-100% white in 2 studies;[9, 27] NR[30]	4	22141	0.58[9] to 0.82[9]	2	3321	33.3% to 37%	74% to 86.4% (85.1-87.7)	MOF ≥9.3%
	Community- based sample, USA[16]	64.2	All men	88.5% white, 8.5% black, 2.9% Mexican-American	1	1498	0.79 (0.74 to 0.84)	1	1498	39% (27-51)	89% (87-91)	MOF risk >=9.3%
	General practice; Australia[40]	78.2	45.1% women	Not reported	1	626	0.68 (0.63 to 0.74)	1	626	89.6 (NR)	35 (NR)	MOF≥6.5%
Gnudi et al., 2005[23] (Age at menarche, weight, years since menopause, previous fracture, weight, fracture in participant's mother, arm help to get up from sitting)	Women requiring a DXA scan at "a center"; Italy	64.3	All women	100% white	1	478	0.74 (0.70 to 0.79)	1	478	95.5%	27.7%	Predicted probability of low BMD at 0.132[c]

© 2018 American Medical Association. All rights reserved.

718 **eTable 60. Characteristics and accuracy of clinical risk assessment tools in identifying osteoporosis (KQ 2a) (continued)**

Instrument (Instrument components)	Clinical and Geographic Setting	Mean Age (Years)	Sex	Race/ Ethnicity	Area under the curve			Sensitivity and specificity				
					No. of Studies	No. of Partici-pants	Pooled AUC (95% CI) or Range of AUCs;[a]	No. of Studies	No. of Partici-pants	Range of Sensitivity	Range of Specificity	Threshold[b]
Mscore[46] 2 models: Age and weight (reduced Mscore) Or Age, weight, gastrectomy, COPD, two or more prior fractures (Mscore)	Clinic-based	60.9 to 68.4	All men	Caucasian	1	197	Age-weight model: 0.81 (0.69 to 0.92) 5-variable model 0.84 (0.74 to 0.95)	1	197	Age-weight model: 85% 5-variable model 88%	Age-weight model: 58% 5-variable model 57%	<9
				African-American	1	134	Age-weight model: 0.99 (0.98 to 1.01); 5-variable model NR	1	134	Age-weight model 93%; 5-variable model NR	Age-weight model: 79% 5-variable model NR	<9
MORES[15, 43, 44] (Age, weight, history of COPD)	1 clinic sample, 2 population-based samples	63 to 70.2	All men	NR	3	4828	0.80 (95% CI, 0.71 to 0.88)	3	4828	66-95%	61-70%	≥6
MOST[31] (QUI, body weight)	Cohort of community-dwelling, ambulatory men; US and Hong Kong	65 and older	All men	71% Caucasian; 29% Chinese	1	4658	US: 0.80 (0.78 to 0.82) Hong Kong: 0.83 (0.80 to 0.86)	NR	NR	NR	NR	NR

719
720

© 2018 American Medical Association. All rights reserved.

721 **eTable 60. Characteristics and accuracy of clinical risk assessment tools in identifying osteoporosis (KQ 2a) (continued)**

Instrument (Instrument components)	Clinical and Geographic Setting	Mean Age (Years)	Sex	Race/ Ethnicity	Area under the curve			Sensitivity and specificity				
					No. of Studies	No. of Partici-pants	Pooled AUC (95% CI) or Range of AUCs;[a]	No. of Studies	No. of Partici-pants	Range of Sensitivity	Range of Specificity	Threshold[b]
NOF guidelines[12, 20, 21, 34] (Age, weight, personal history of fracture with minimal trauma > 40 years, family history of fracture, current cigarette smoking)	Majority of studies general population or general practice; USA, Canada, Italy	57.3 to 69.2	All women	Predominantly white	2	1520	0.60	2	2567	96-100%	10-18%	≥1
ORAI[12-14, 17, 18, 20-22, 24-27, 33, 34, 37, 42, 137] (Age, weight in pounds, current estrogen use)	Half of the studies conducted in general practice or population settings; USA. Australia, Belgium, Canada, Denmark, England. Italy. Singapore. Spain	50.5 to 70.5	All women	White participants in majority of studies	10	16,780	0.65 (0.60 to 0.71)	9	7830	50-100%	10-75%	≥9
OSIRIS[2, 18, 26, 27, 33, 42] (Age, weight, HRT use, history of low trauma fracture)	All clinic-based, all in Europe	54.1 to 61.5	All women	Predominantly Causasian	5	649	0.68 (0.64 to 0.72)	1	4035	64%	69%	<1
OST[8, 31, 32, 41, 45, 46] (Age and weight)	4 clinic-based, 2 community-based; 5 in US and 1 in Portugal	64 to 68	All men	Predominantly Caucasian	6	7798	0.76 (0.71 to 0.80)	5	5,366	61.8% to 87.6%	36.1 to 74%	<2
OST[2, 13, 18-22, 25-27, 30, 33, 35, 36, 42, 137] (Age and weight)	9 clinic-based and 6 community based; 3 in US, 4 in Canada, 8 in Northern/Western Europe	51 to 62	All women	Predominantly Caucasian	13	44323	0.65 (0.60 to 0.69)	7	11486	69% to 95.3%	NR	<2
					11	42802	without outlier,[20, 21] 0.71 (0.70 to 0.72)	11	42802	NR	34% to 71%	<2
OST[40] (Age and weight)	Clinic-based, Australia	78	45.1% men	Predominantly Caucasian	1	626	0.76 (0.71 to 0.82)	1	626	Any site	39.9%	≤0

722
723

© 2018 American Medical Association. All rights reserved.

eTable 60. Characteristics and accuracy of clinical risk assessment tools in identifying osteoporosis (KQ 2a) (continued)

					Area under the curve			Sensitivity and specificity				
Instrument (Instrument components)	**Clinical and Geographic Setting**	**Mean Age (Years)**	**Sex**	**Race/ Ethnicity**	**No. of Studies**	**No. of Partici-pants**	**Pooled AUC (95% CI) or Range of AUCs;[a]**	**No. of Studies**	**No. of Partici-pants**	**Range of Sensitivity**	**Range of Specificity**	**Threshold[b]**
OSTA[29, 32, 38] (Age and weight)	Community-based, Hong Kong and S. Korea	63.4 to 54	All men	Asian	2	1,911	0.627 to 0.72	2	1,911	Varies by study, no common cutoff	Varies by study, no common cutoff	No common cutoff
OSTA[17, 21, 28, 37, 38, 47] (Age and weight)	1 clinic-based and 4 community-based studies; Australia, Singapore, Hong Kong, South Korea	59.1 to 70.5	All women	Caucasian and Asian	4	2962	0.76 (0.63 to 0.90)	5	3,414	41% to 97%	24% to 67.1%	≤-1
SCORE[11, 12, 14, 17-19, 22, 24, 26, 27, 34, 42, 137] (Age, weight, and estrogen replacement therapy, the SCORE instrument includes race/ethnicity, history of rheumatoid arthritis, and history of nontraumatic fractures after age 45)	4 clinic-based, 7 community-based US: 4; UK: 2; Spain: 1: Singapore: 1; Belgium: 1; Denmark: 1; Canada: 1	57.7 to 69.2	All women	Predominantly white	8	15,362	0.70, (0.69 to 0.71)	6	7455	54% to 100%	17.9% to 72%	≥6

© 2018 American Medical Association. All rights reserved.

eTable 60. Characteristics and accuracy of clinical risk assessment tools in identifying osteoporosis (KQ 2a) (continued)

Instrument (Instrument components)	Clinical and Geographic Setting	Mean Age (Years)	Sex	Race/ Ethnicity	Area under the curve			Sensitivity and specificity				
					No. of Studies	No. of Partici-pants	Pooled AUC (95% CI) or Range of AUCs;[a]	No. of Studies	No. of Partici-pants	Range of Sensitivity	Range of Specificity	Threshold[b]
SOF[11] (Prior fracture after age 50; age 60-64 with t-score < than -2.5 or age 65 or older with z-score < than -0.43; and 5 or more risk factors (first-degree relative with hip fracture, current weight less than at age 25, dementia, using corticosteroids or seizure medication or benzodiazepines, had a fracture age 50+, not taking HRT, on feet <4 h/day, heart rate >80 beats/min, was >5'7 at age 25, 80+ years old; subtract 1 point each for race (African American); walk for exercise; can rise from chair without arms)	OPRA study, Group Health participant; US	69.3	All women	93.5% White	1	416	0.54 (SE 0.03); 1; 416	1	416	32.6 (26.6, 38.6)	76.0 (63.5, 88.6)	≥ 5
SOFSURF[18, 22, 37] (Age, weight, smoking and history of postmenopausal fracture)	Population-based cohort; Dubbo, Australia Scanning clinics; UK	59.7 to 70.5	All women	Mostly white	1	208	0.717 (0.777 to 0.670)[18] NR in 2 studies;[22, 37]			Varies by study, no common cutoff	Varies by study, no common cutoff	No common cutoff

[a] Presented for any site when available (femoral neck, lumbar spine, total hip); if not available, presented for femoral neck
[b] Sensitivity, specificity, NPV, and PPV presented for the most commonly reported threshold across studies
[c] Study presents multiple predicted probabilities of low BMD; the study notes that the threshold offered the highest number of DXA-deferred cases and the lowest number of low-BMD missed cases.
[d] Studies present results for three different sites of BMD measurement: total hip,[43] total hip or femoral neck,[15] or thoracic vertebra, lumbar vertebra, arms, ribs, pelvis, or legs[44]
[e] The African-American sample includes data from 95 new participants and 39 participants from development cohort and is therefore not a pure validation cohort

© 2018 American Medical Association. All rights reserved.

Abbreviations: ABONE = assessing age, body size, and estrogen use; AMMEB = Age, Years after Menopause, Age at Menarche, Body Mass Index; AUC= area under the curve; BMD= bone mineral density; BMI = body mass index; CI = confidence interval; COPD = Chronic Obstructive Pulmonary Disease; DOEScore = Dubbo Osteoporosis Epidemiology Score; DXA= Dual-energy X-ray absorptiometry; FN= Femoral neck; FRAX = Fracture Risk Assessment tool; HRT = hormone replacement therapy; MOF= Melton Osteoporotic Fracture study; MORE = Multiple Outcomes of Raloxifene Trial; MOST = Male Osteoporosis Screening Tool; NOF = National Osteoporosis Foundation; NR = not reported; OPRA= osteoporosis population-based risk assessment ORAI = Osteoporosis Risk Assessment Instrument; OSIRIS = Osteoporosis Index of Risk; OST = osteoporosis self-assessment tool; OSTA = Osteoporosis Self-assessment Tool for Asians; QUI = ultrasound index; SCORE = Simple Calculated Osteoporosis Risk Estimation Tool; SE = standard error; SOF = Study of Osteoporotic Fractures; SOFSURF = Study of Osteoporotic Fractures Simple Useful Risk Factors ;UK= United Kingdom; US = United States; USA= United States of America.

© 2018 American Medical Association. All rights reserved.

eTable 61. Characteristics and accuracy of bone measurement tests in identifying osteoporosis (KQ 2a)

Imaging Test	Site of Test	Number of Studies	Number of Participants	Sex	Mean Age Range (Years)	Gold Standard Test	Site of Gold Standard	Summary of Accuracy (AUC; Range of AUCs, Pooled AUC (95% CI) >1 study)
QUS	Calcaneus	7[18, 26, 28, 35, 48-50]	1,969	Women	59–63	DXA ≤-2.5	Lumbar spine, femoral, or total hip	Range 0.69 to 0.898 Pooled estimate: 0.77 (95% CI, 0.72 to 0.82)
		3[29, 31, 45]	5,142	Men	61–63	DXA ≤-2.5	Lumbar spine, femoral, or total hip	Range 0.696 to 0.93 Pooled estimate: 0.80 (95% CI, 0.67 to 0.94)
Peripheral DXA	Calcaneus	2[26, 27]	712	Women	61 (SD ranges from 4 to 8)	DXA	Lumbar spine, femoral, or total hip	Range0.67 to 0.803 (variance NR)
DXR	Nondominant metacarpals	1[48]	221	Women	61 (range 50–75)	DXA	Lumbar spine or total hip	0.84 (95% CI, 0.79 to 0.89)
RA	Nondominant phalanges	1[48]	221	Women	61 (range 50–75)	DXA	Lumbar spine or total hip	0.80 (95% CI, 0.74 to 0.85)

Abbreviations: AUC = area under the curve; CI = confidence interval; DXA = dual energy X-ray absorptiometry; DXR = digital X-ray radiogrammetry; NR = not reported; QUS = quantitative ultrasound; RA = radiographic absorptiometry; SD = standard deviation; SE = standard error.

© 2018 American Medical Association. All rights reserved.

eTable 62. Using repeat BMD testing to predict fracture risk (KQ 2b)

Study	Study Cohort*, Country	Inclusion / Exclusion Criteria	Mean Length of Followup, Years (Range)	N	Participant Characteristics	Bone Measurement Test	Fracture Site	AUC for Baseline BMD (95% CI)	AUC for BMD% Change (95% CI)	AUC for BMD Baseline and% Change (95% CI)
Berry, 2013[54]	Framingham Osteoporosis Study, USA	Included participants with at least two BMD measurements. Excluded those with fracture prior to second test.	3.7 (2.4 to 6.0)	802	Mean age: 74.8 (SD 4.5) Percent women: 61	DXA, BMD	Hip fracture[a]	0.71 (0.65 to 0.78)	0.68 (0.62 to 0.75)	0.72 (0.66 to 0.79)
							MOF fracture[a]	0.74 (0.69 to 0.79)	0.71 (0.66 to 0.76)	0.74 (0.69 to 0.79)
Hillier, 2007[59]	Study of Osteoporotic Fractures, USA	Included participants with at least two BMD measurements. Excluded those with fracture prior to second test.	8.0 (6.3 to 9.8)	4,124	Mean age: 74 (SD 4) Percent women: 100	DXA, BMD	Hip fracture[b]	0.73 (CI, NR)	0.68 (CI, NR)	0.74 (CI, NR)
							Nonspine fracture[b]	0.65 (CI ,NR)	0.61 (CI, NR)	0.65 (CI, NR)
							Spine fracture[b]	0.67 (CI, NR)	0.62 (CI, NR)	0.68 (CI, NR)

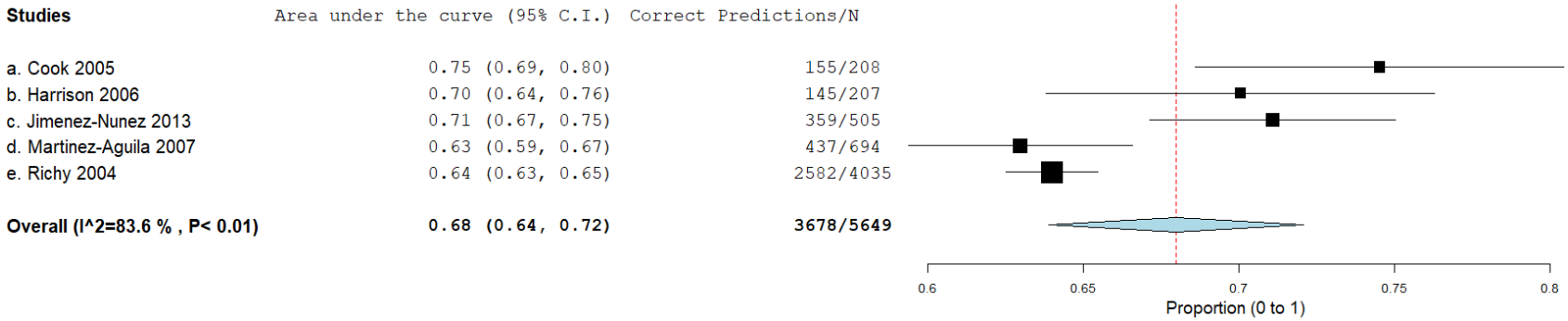
[a] Adjusted for age, sex, BMI, weight loss, and history of fracture measured at the time of the second BMD.

[b] Adjusted for age and weight change.

Abbreviations: AUC=area under the curve; BMD=bone mineral density; BMI=body mass index; CI=confidence interval; DXA=dual energy X-ray absorptiometry; MOF=major osteoporotic fracture defined as fractures of the proximal femur, distal radius, proximal humerus, and clinical vertebral fractures; N=number; NR=not reported; SD=standard deviation; USA=United States of America.

© 2018 American Medical Association. All rights reserved.

eFigure 1. Osteoporosis Risk Assessment Instrument (ORAI) in women (KQ 2a)



[a] Cook RB, Collins D, Tucker J, et al Comparison of questionnaire and quantitative ultrasound techniques as screening tools for DXA. *Osteoporos Int*. 2005 Dec;16(12):1565-75. doi: 10.1007/s00198-005-1864-x [doi]. PMID: 15883661.]

[b] Harrison EJ, Adams JE. Application of a triage approach to peripheral bone densitometry reduces the requirement for central DXA but is not cost effective. *Calcif Tissue Int*. 2006 Oct;79(4):199-206. doi: 10.1007/s00223-005-0302-6. PMID: 16969598.

[c] Jimenez-Nunez FG, Manrique-Arija S, Urena-Garnica I, et al Reducing the need for central dual-energy X-ray absorptiometry in postmenopausal women: efficacy of a clinical algorithm including peripheral densitometry. *Calcif Tissue Int*. 2013 Jul;93(1):62-8. doi: 10.1007/s00223-013-9728-4 [doi]. PMID: 23608922.

[d] Martinez-Aguila D, Gomez-Vaquero C, Rozadilla A, et al Decision rules for selecting women for bone mineral density testing: application in postmenopausal women referred to a bone densitometry unit. *J Rheumatol*. 2007 Jun;34(6):1307-12. PMID: 17552058.

[e] Richy F, Gourlay M, Ross PD, et al Validation and comparative evaluation of the osteoporosis self-assessment tool (OST) in a Caucasian population from Belgium. *QJM*. 2004 Jan;97(1):39-46. PMID: 14702510.

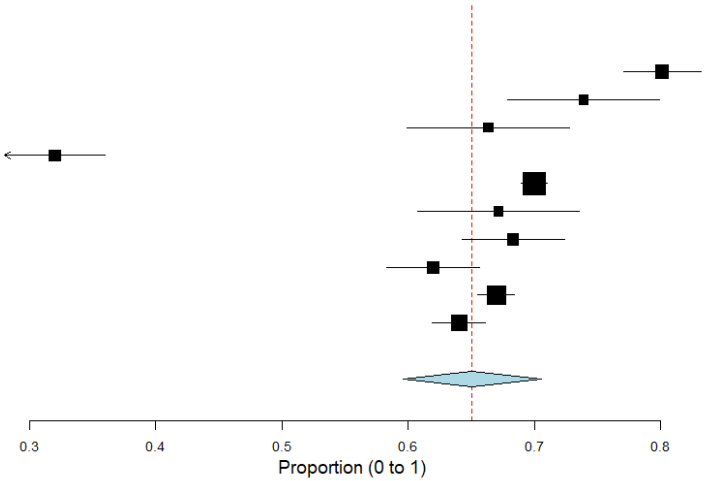
Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 2. Osteoporosis Index of Risk (OSIRIS) in women (KQ 2a)

Studies	Area under the curve (95% C.I.)	Correct predictions/N
a. Cadarette 2004	0.80 (0.77, 0.83)	516/644
b. Cass 2006	0.74 (0.68, 0.80)	150/203
c. Cook 2005	0.66 (0.60, 0.73)	138/208
d. D'Amelio 2005	0.32 (0.28, 0.36)	168/525
e. Gourlay 2008	0.70 (0.69, 0.71)	5445/7779
f. Harrison 2006	0.67 (0.61, 0.74)	139/207
g. Jimenez-Nunez 2013	0.68 (0.64, 0.72)	345/505
h. Martinez-Aguila 2007	0.62 (0.58, 0.66)	412/665
i. Richy 2004	0.67 (0.66, 0.68)	2703/4035
j. Rud 2005	0.64 (0.62, 0.66)	1286/2009
Overall (I^2=97.84 % , P< 0.01)	**0.65 (0.60, 0.71)**	**11302/16780**



[a] Cadarette SM, McIsaac WJ, Hawker GA, et al The validity of decision rules for selecting women with primary osteoporosis for bone mineral density testing. *Osteoporos Int*. 2004 May;15(5):361-6. doi: 10.1007/s00198-003-1552-7. PMID: 14730421.

[b] Cass AR, Shepherd AJ, Carlson CA. Osteoporosis risk assessment and ethnicity: validation and comparison of 2 clinical risk stratification instruments. *J Gen Intern Med*. 2006 Jun;21(6):630-5. doi: 10.1111/j.1525-1497.2006.00459.x. PMID: 16808748.

[c] Cook RB, Collins D, Tucker J, et al Comparison of questionnaire and quantitative ultrasound techniques as screening tools for DXA. *Osteoporos Int*. 2005 Dec;16(12):1565-75. doi: 10.1007/s00198-005-1864-x [doi]. PMID: 15883661.]

[d] D'Amelio P, Tamone C, Pluviano F, et al Effects of lifestyle and risk factors on bone mineral density in a cohort of Italian women: suggestion for a new decision rule. Calcif Tissue Int. 2005 Aug;77(2):72-8. doi: 10.1007/s00223-004-0253-3. PMID: 16059776.

[e] Gourlay ML, Powers JM, Lui LY, et al Clinical performance of osteoporosis risk assessment tools in women aged 67 years and older. *Osteoporos Int*. 2008 Aug;19(8):1175-83. doi: 10.1007/s00198-007-0555-1. PMID: 18219434.

[f] Harrison EJ, Adams JE. Application of a triage approach to peripheral bone densitometry reduces the requirement for central DXA but is not cost effective. *Calcif Tissue Int*. 2006 Oct;79(4):199-206. doi: 10.1007/s00223-005-0302-6. PMID: 16969598.

[g] Jimenez-Nunez FG, Manrique-Arija S, Urena-Garnica I, et al Reducing the need for central dual-energy X-ray absorptiometry in postmenopausal women: efficacy of a clinical algorithm including peripheral densitometry. *Calcif Tissue Int*. 2013 Jul;93(1):62-8. doi: 10.1007/s00223-013-9728-4 [doi]. PMID: 23608922.

[h] Martinez-Aguila D, Gomez-Vaquero C, Rozadilla A, et al Decision rules for selecting women for bone mineral density testing: application in postmenopausal women referred to a bone densitometry unit. *J Rheumatol*. 2007 Jun;34(6):1307-12. PMID: 17552058.

[i] Richy F, Gourlay M, Ross PD, et al Validation and comparative evaluation of the osteoporosis self-assessment tool (OST) in a Caucasian population from Belgium. *QJM*. 2004 Jan;97(1):39-46. PMID: 14702510.

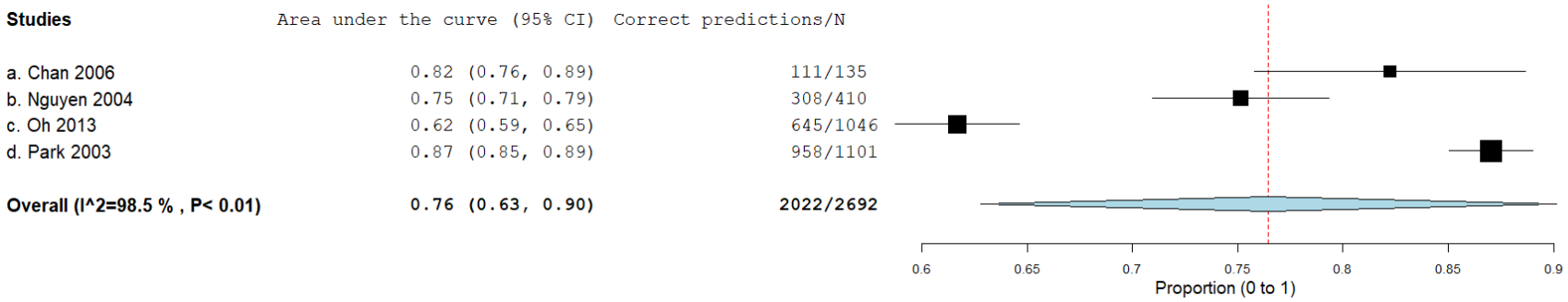
[j] Rud B, Jensen JE, Mosekilde L, et al Performance of four clinical screening tools to select peri- and early postmenopausal women for dual X-ray absorptiometry. Osteoporos Int. 2005 Jul;16(7):764-72. doi: 10.1007/s00198-004-1748-5. PMID: 15986263.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 3. Osteoporosis Self-Assessment Tool in Asians (OSTA) in women (KQ 2a)



[a] Chan SP, Teo CC, Ng SA, et al Validation of various osteoporosis risk indices in elderly Chinese females in Singapore. *Osteoporos Int*. 2006;17(8):1182-8. doi: 10.1007/s00198-005-0051-4 [doi]. PMID: 16699739.

[b] Nguyen TV, Center JR, Pocock NA, et al Limited utility of clinical indices for the prediction of symptomatic fracture risk in postmenopausal women. *Osteoporos Int*. 2004 Jan;15(1):49-55. doi: 10.1007/s00198-003-1511-3. PMID: 14593453.

[c] Oh SM, Nam BH, Rhee Y, et al Development and validation of osteoporosis risk-assessment model for Korean postmenopausal women. *J Bone Miner Metab*. 2013 Jul;31(4):423-32. doi: 10.1007/s00774-013-0426-0 [doi]. PMID: 23420298.

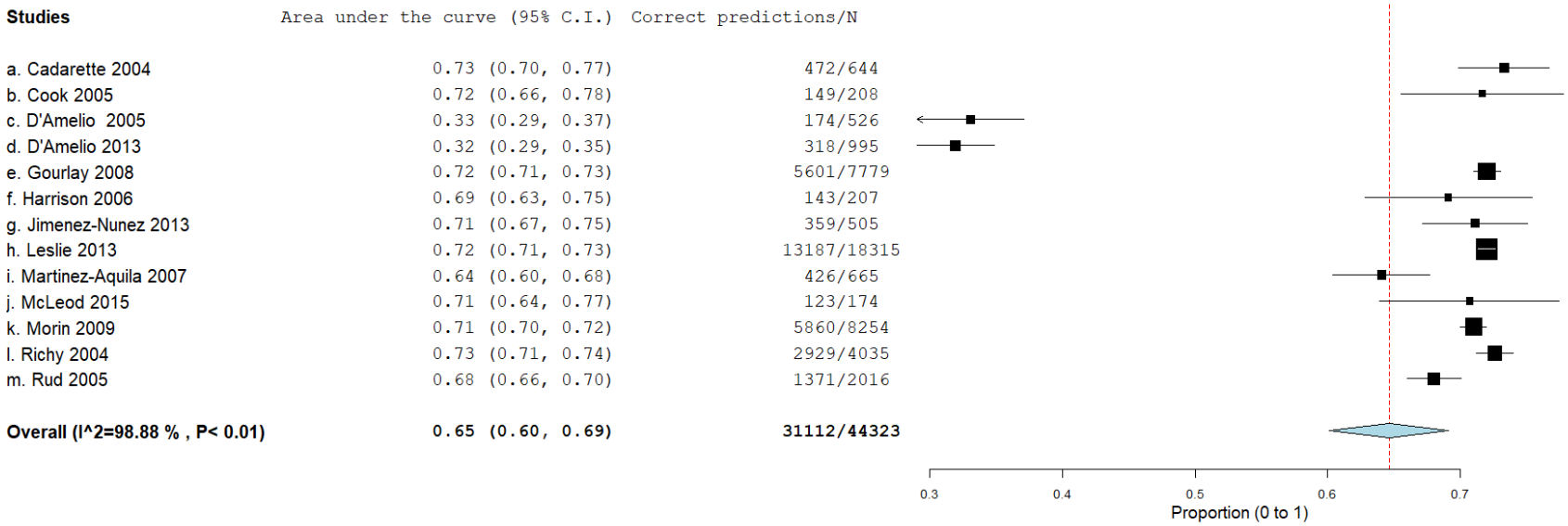
[d] Park HM, Sedrine WB, Reginster JY, et al Korean experience with the OSTA risk index for osteoporosis: a validation study. *J Clin Densitom*. 2003 Fall;6(3):247-50. PMID: 14514994.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 4. Osteoporosis Self-Assessment Tool (OST) in women (KQ 2a)



[a] Cadarette SM, McIsaac WJ, Hawker GA, et al The validity of decision rules for selecting women with primary osteoporosis for bone mineral density testing. *Osteoporos Int*. 2004 May;15(5):361-6. doi: 10.1007/s00198-003-1552-7. PMID: 14730421.

[b] Cook RB, Collins D, Tucker J, et al Comparison of questionnaire and quantitative ultrasound techniques as screening tools for DXA. *Osteoporos Int*. 2005 Dec;16(12):1565-75. doi: 10.1007/s00198-005-1864-x [doi]. PMID: 15883661.]

[c] D'Amelio P, Tamone C, Pluviano F, et al Effects of lifestyle and risk factors on bone mineral density in a cohort of Italian women: suggestion for a new decision rule. Calcif Tissue Int. 2005 Aug;77(2):72-8. doi: 10.1007/s00223-004-0253-3. PMID: 16059776.

[d] D'Amelio P, Spertino E, Martino F, et al Prevalence of postmenopausal osteoporosis in Italy and validation of decision rules for referring women for bone densitometry. Calcif Tissue Int. 2013;92(5):437-43.

[e] Gourlay ML, Powers JM, Lui LY, et al Clinical performance of osteoporosis risk assessment tools in women aged 67 years and older. *Osteoporos Int*. 2008 Aug;19(8):1175-83. doi: 10.1007/s00198-007-0555-1. PMID: 18219434.

[f] Harrison EJ, Adams JE. Application of a triage approach to peripheral bone densitometry reduces the requirement for central DXA but is not cost effective. *Calcif Tissue Int*. 2006 Oct;79(4):199-206. doi: 10.1007/s00223-005-0302-6. PMID: 16969598.

[g] Jimenez-Nunez FG, Manrique-Arija S, Urena-Garnica I, et al Reducing the need for central dual-energy X-ray absorptiometry in postmenopausal women: efficacy of a clinical algorithm including peripheral densitometry. *Calcif Tissue Int*. 2013 Jul;93(1):62-8. doi: 10.1007/s00223-013-9728-4 [doi]. PMID: 23608922.

[h] Leslie WD, Lix LM, Johansson H, et al Selection of women aged 50-64 yr for bone density measurement. J Clin Densitom. 2013 Oct-Dec;16(4):570-8. doi: 10.1016/j.jocd.2013.01.004. PMID: 23452870.

[i] Martinez-Aguila D, Gomez-Vaquero C, Rozadilla A, et al Decision rules for selecting women for bone mineral density testing: application in postmenopausal women referred to a bone densitometry unit. *J Rheumatol*. 2007 Jun;34(6):1307-12. PMID: 17552058.

[j] McLeod KM, Johnson S, Rasali D, et al Discriminatory performance of the calcaneal quantitative ultrasound and osteoporosis self-assessment tool to select older women for dual-energy x-ray absorptiometry. J Clin Densitom. 2015 Apr-Jun;18(2):157-64. doi: 10.1016/j.jocd.2015.02.006. PMID: 25937306.

© 2018 American Medical Association. All rights reserved.

[k] Morin S, Tsang JF, Leslie WD. Weight and body mass index predict bone mineral density and fractures in women aged 40 to 59 years. Osteoporos Int. 2009 Mar;20(3):363-70. doi: 10.1007/s00198-008-0688-x [doi]. PMID: 18633665.
[l] Richy F, Gourlay M, Ross PD, et al Validation and comparative evaluation of the osteoporosis self-assessment tool (OST) in a Caucasian population from Belgium. *QJM*. 2004 Jan;97(1):39-46. PMID: 14702510.
[m] Rud B, Jensen JE, Mosekilde L, et al Performance of four clinical screening tools to select peri- and early postmenopausal women for dual X-ray absorptiometry. Osteoporos Int. 2005 Jul;16(7):764-72. doi: 10.1007/s00198-004-1748-5. PMID: 15986263.

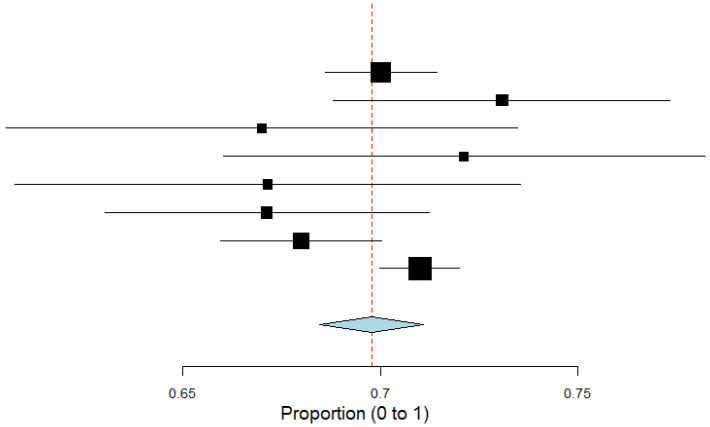
Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 5. Simple Calculated Osteoporosis Risk Estimation (SCORE) in women (KQ 2a)

Studies	Area under the curve (95% C.I.)	Correct predictions/N
a. Ben Sedrine 2001	0.70 (0.69, 0.71)	2825/4035
b. Brenneman 2003	0.73 (0.69, 0.77)	304/416
c. Cass 2006	0.67 (0.61, 0.73)	136/203
d. Cook 2005	0.72 (0.66, 0.78)	150/208
e. Harrison 2006	0.67 (0.61, 0.74)	139/207
f. Jimenez-Nunez 2013	0.67 (0.63, 0.71)	339/505
g. Rud 2005	0.68 (0.66, 0.70)	1366/2009
h. Gourlay 2008	0.71 (0.70, 0.72)	5523/7779
Overall (I^2=46.45 % , P=0.07)	**0.70 (0.68, 0.71)**	**10782/15362**



[a] Ben Sedrine W, Devogelaer JP, Kaufman JM, et al Evaluation of the simple calculated osteoporosis risk estimation (SCORE) in a sample of white women from Belgium. Bone. 2001 Oct;29(4):374-80. PMID: 11595621.

[b] Brenneman SK, Lacroix AZ, Buist DS, et al Evaluation of decision rules to identify postmenopausal women for intervention related to osteoporosis. Dis Manag. 2003 Fall;6(3):159-68. doi: 10.1089/109350703322425509. PMID: 14570384.

[c] Cass AR, Shepherd AJ, Carlson CA. Osteoporosis risk assessment and ethnicity: validation and comparison of 2 clinical risk stratification instruments. J Gen Intern Med. 2006 Jun;21(6):630-5. doi: 10.1111/j.1525-1497.2006.00459.x. PMID: 16808748.

[d] Cook RB, Collins D, Tucker J, et al Comparison of questionnaire and quantitative ultrasound techniques as screening tools for DXA. Osteoporos Int. 2005 Dec;16(12):1565-75. doi: 10.1007/s00198-005-1864-x [doi]. PMID: 15883661.]

[e] Harrison EJ, Adams JE. Application of a triage approach to peripheral bone densitometry reduces the requirement for central DXA but is not cost effective. Calcif Tissue Int. 2006 Oct;79(4):199-206. doi: 10.1007/s00223-005-0302-6. PMID: 16969598.

[f] Jimenez-Nunez FG, Manrique-Arija S, Urena-Garnica I, et al Reducing the need for central dual-energy X-ray absorptiometry in postmenopausal women: efficacy of a clinical algorithm including peripheral densitometry. Calcif Tissue Int. 2013 Jul;93(1):62-8. doi: 10.1007/s00223-013-9728-4 [doi]. PMID: 23608922.

[g] Rud B, Jensen JE, Mosekilde L, et al Performance of four clinical screening tools to select peri- and early postmenopausal women for dual X-ray absorptiometry. Osteoporos Int. 2005 Jul;16(7):764-72. doi: 10.1007/s00198-004-1748-5. PMID: 15986263.

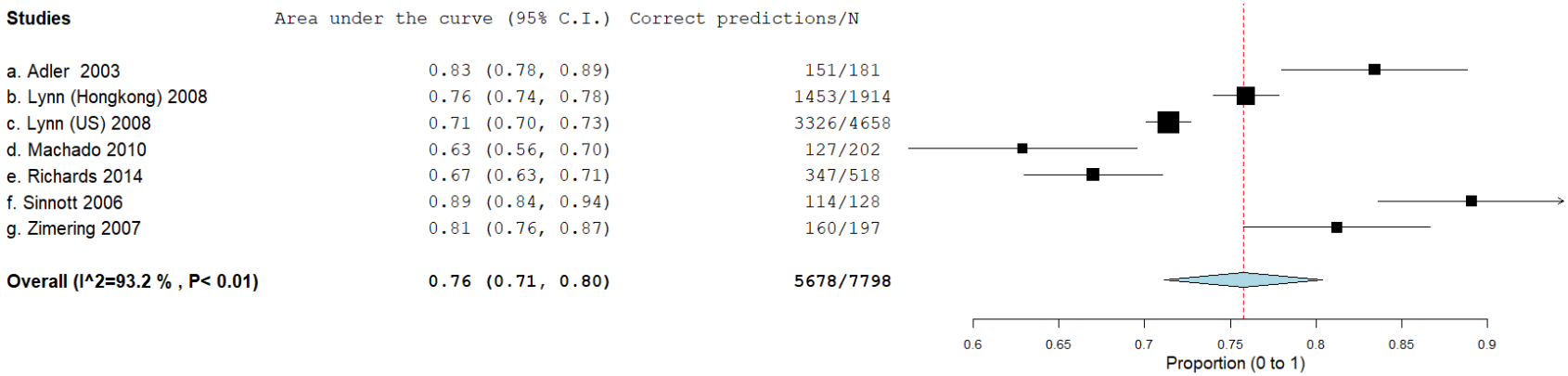
[h] Gourlay ML, Powers JM, Lui LY, et al Clinical performance of osteoporosis risk assessment tools in women aged 67 years and older. Osteoporos Int. 2008 Aug;19(8):1175-83. doi: 10.1007/s00198-007-0555-1. PMID: 18219434.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 6. OST in men (KQ 2a)



[a] Adler RA, Tran MT, Petkov VI. Performance of the Osteoporosis Self-assessment Screening Tool for osteoporosis in American men. *Mayo Clin Proc*. 2003 Jun;78(6):723-7. doi: 10.4065/78.6.723. PMID: 12934782.

[b] Lynn HS, Woo J, Leung PC, et al An evaluation of osteoporosis screening tools for the osteoporotic fractures in men (MrOS) study. *Osteoporos Int*. 2008 Jul;19(7):1087-92. doi: 10.1007/s00198-007-0553-3. PMID: 18239959.

[c] Machado P, Coutinho M, da Silva JA. Selecting men for bone densitometry: performance of osteoporosis risk assessment tools in Portuguese men. *Osteoporos Int*. 2010 Jun;21(6):977-83. doi: 10.1007/s00198-009-1036-5 [doi]. PMID: 19727909.

[d] 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.

[e] 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. doi: 10.1007/s00198-005-0034-5. PMID: 16523248.

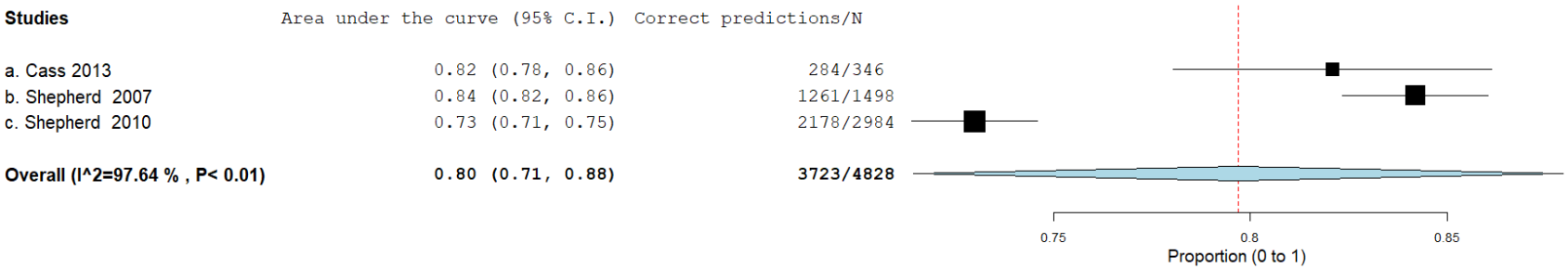
[f] Zimering MB, Shin JJ, Shah J, et al Validation of a novel risk estimation tool for predicting low bone density in Caucasian and African American men veterans. *J Clin Densitom*. 2007 Jul-Sep;10(3):289-97. doi: 10.1016/j.jocd.2007.03.001. PMID: 17459748.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 7. Male Osteoporosis Risk Estimation Score (MORES) in men (KQ 2a)



[a] Cass AR, Shepherd AJ. Validation of the Male Osteoporosis Risk Estimation Score (MORES) in a primary care setting. *J Am Board Fam Med*. 2013 Jul-Aug;26(4):436-44. doi: 10.3122/jabfm.2013.04.120182. PMID: 23833159.

[b] Shepherd AJ, Cass AR, Carlson CA, et al Development and internal validation of the male osteoporosis risk estimation score. *Ann Fam Med*. 2007 Nov-Dec;5(6):540-6. doi: 10.1370/afm.753. PMID: 18025492.

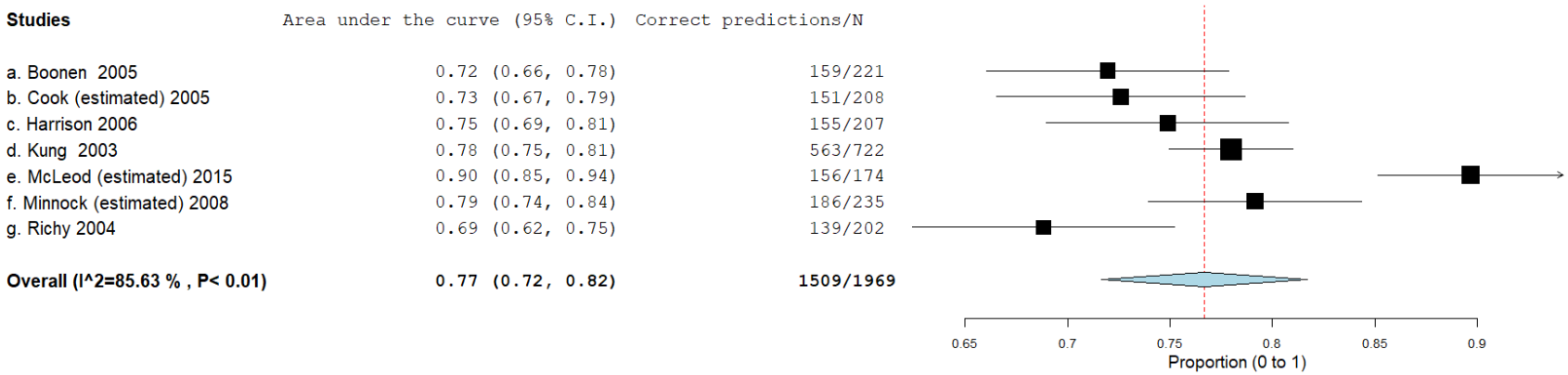
[c] Shepherd AJ, Cass AR, Ray L. Determining risk of vertebral osteoporosis in men: validation of the male osteoporosis risk estimation score. *J Am Board Fam Med*. 2010 Mar-Apr;23(2):186-94. doi: 10.3122/jabfm.2010.02.090027. PMID: 20207929.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 8. Quantitative ultrasound for screening osteoporosis for women (KQ 2a)



[a] Boonen S, Nijs J, Borghs H, et al. Identifying postmenopausal women with osteoporosis by calcaneal ultrasound, metacarpal digital X-ray radiogrammetry and phalangeal radiographic absorptiometry: a comparative study. Osteoporos Int. 2005 Jan;16(1):93-100. doi: 10.1007/s00198-004-1660-z [doi]. PMID: 15197540.

[b] Cook RB, Collins D, Tucker J, et al. Comparison of questionnaire and quantitative ultrasound techniques as screening tools for DXA. Osteoporos Int. 2005 Dec;16(12):1565-75. doi: 10.1007/s00198-005-1864-x [doi]. PMID: 15883661

[c] Harrison EJ, Adams JE. Application of a triage approach to peripheral bone densitometry reduces the requirement for central DXA but is not cost effective. Calcif Tissue Int. 2006 Oct;79(4):199-206. doi: 10.1007/s00223-005-0302-6. PMID: 16969598.

[d] Kung AW, Ho AY, Sedrine WB, et al. Comparison of a simple clinical risk index and quantitative bone ultrasound for identifying women at increased risk of osteoporosis. Osteoporos Int. 2003 Sep;14(9):716-21. doi: 10.1007/s00198-003-1428-x [doi]. PMID: 12897978.

[e] McLeod KM, Johnson S, Rasali D, et al. Discriminatory performance of the calcaneal quantitative ultrasound and osteoporosis self-assessment tool to select older women for dual-energy x-ray absorptiometry. J Clin Densitom. 2015 Apr-Jun;18(2):157-64. doi: 10.1016/j.jocd.2015.02.006. PMID: 25937306.

[f] Minnock E, Cook R, Collins D, et al. Using risk factors and quantitative ultrasound to identify postmenopausal caucasian women at risk of osteoporosis. J Clin Densitom. 2008 Oct-Dec;11(4):485-93. doi: 10.1016/j.jocd.2008.04.002. PMID: 18539491.

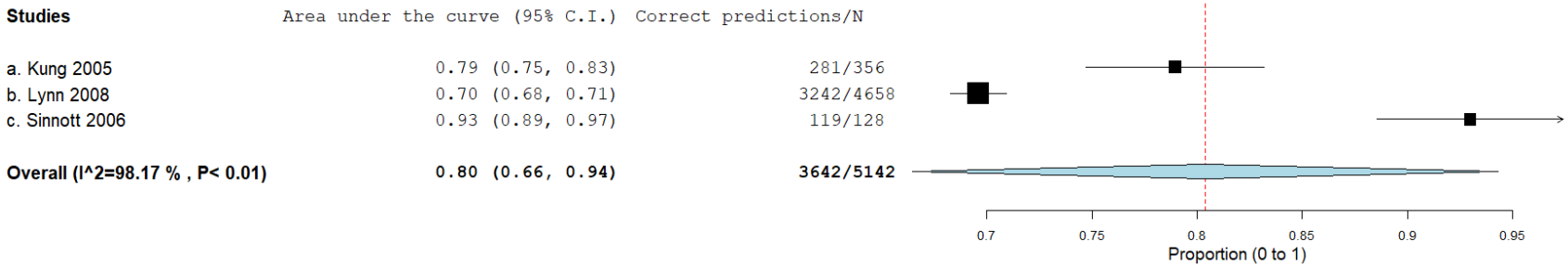
[g] Richy F, Deceulaer F, Ethgen O, et al. Development and validation of the ORACLE score to predict risk of osteoporosis. Mayo Clin Proc. 2004 Nov;79(11):1402-8. doi: 10.4065/79.11.1402. PMID: 15544019.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 9. Quantitative ultrasound for screening osteoporosis for men (KQ 2a)



[a] Kung AW, Ho AY, Ross PD, et al. Development of a clinical assessment tool in identifying Asian men with low bone mineral density and comparison of its usefulness to quantitative bone ultrasound. Osteoporos Int. 2005 Jul;16(7):849-55. doi: 10.1007/s00198-004-1778-z [doi]. PMID: 15611839.

[a] Lynn HS, Woo J, Leung PC, et al. An evaluation of osteoporosis screening tools for the osteoporotic fractures in men (MrOS) study. Osteoporos Int. 2008 Jul;19(7):1087-92. doi: 10.1007/s00198-007-0553-3. PMID: 18239959.

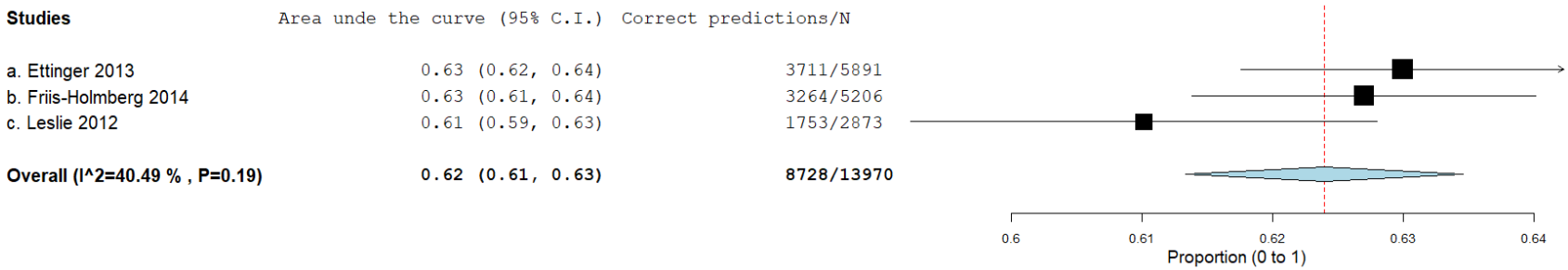
[a] Sinnott B, Kukreja S, Barengolts E. Utility of screening tools for the prediction of low bone mass in African American men. Osteoporos Int. 2006;17(5):684-92. doi: 10.1007/s00198-005-0034-5. PMID: 16523248.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 10. FRAX without bone mineral density testing for predicting major osteoporotic fractures in men (KQ 2a)



[a] 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 Apr;24(4):1185-93. doi: 10.1007/s00198-012-2215-3 [doi]. PMID: 23179575.

[b] 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 Jan-Mar;17(1):7-15. doi: 10.1016/j.jocd.2013.03.014. PMID: 23623379.

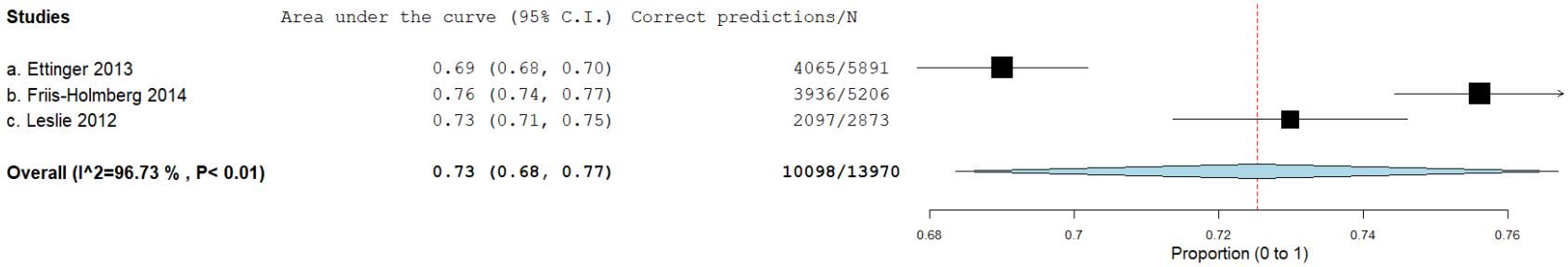
[c] Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. Osteoporos Int. 2012 Jan;23(1):75-85. doi: 10.1007/s00198-011-1747-2. PMID: 21850546

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 11. FRAX without bone mineral density testing for predicting hip fractures in men (KQ 2a)



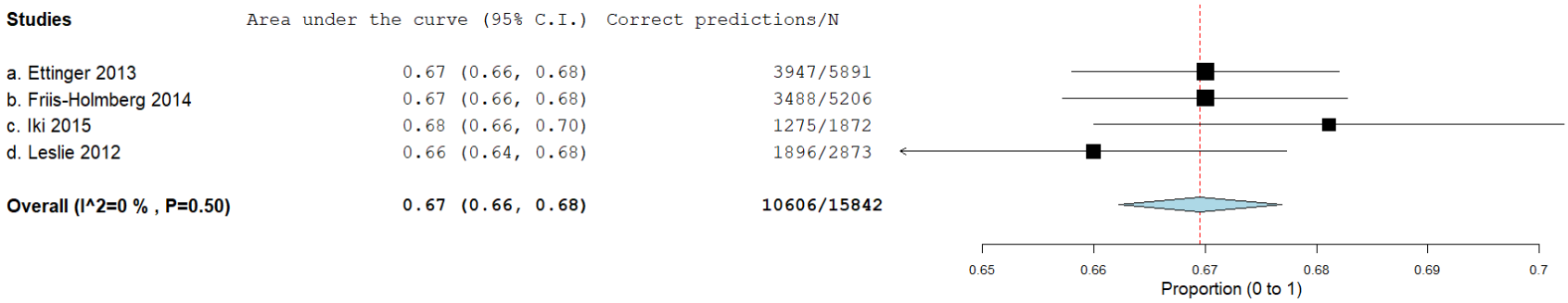
[a] Ettinger B, Ensrud KE, Blackwell T, et al Performance of FRAX in a cohort of community-dwelling, ambulatory older men: the Osteoporotic Fractures in Men (MrOS) study. Osteoporos Int. 2013 Apr;24(4):1185-93. doi: 10.1007/s00198-012-2215-3 [doi]. PMID: 23179575.
[a] Friis-Holmberg T, Rubin KH, Brixen 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 Jan-Mar;17(1):7-15. doi: 10.1016/j.jocd.2013.03.014. PMID: 23623379.
[a] Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. Osteoporos Int. 2012 Jan;23(1):75-85. doi: 10.1007/s00198-011-1747-2. PMID: 21850546

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 12. FRAX with bone mineral density testing for predicting major osteoporotic fractures in men (KQ 2a)



[a] 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 Apr;24(4):1185-93. doi: 10.1007/s00198-012-2215-3 [doi]. PMID: 23179575.

[b] 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 Jan-Mar;17(1):7-15. doi: 10.1016/j.jocd.2013.03.014. PMID: 23623379.

[c] Iki M, Fujita Y, Tamaki J, et al. Trabecular bone score may improve FRAX(R) prediction accuracy for major osteoporotic fractures in elderly Japanese men: the Fujiwara-kyo Osteoporosis Risk in Men (FORMEN) Cohort Study. Osteoporos Int. 2015;26(6):1841-8. doi: 10.1007/s00198-015-3092-3

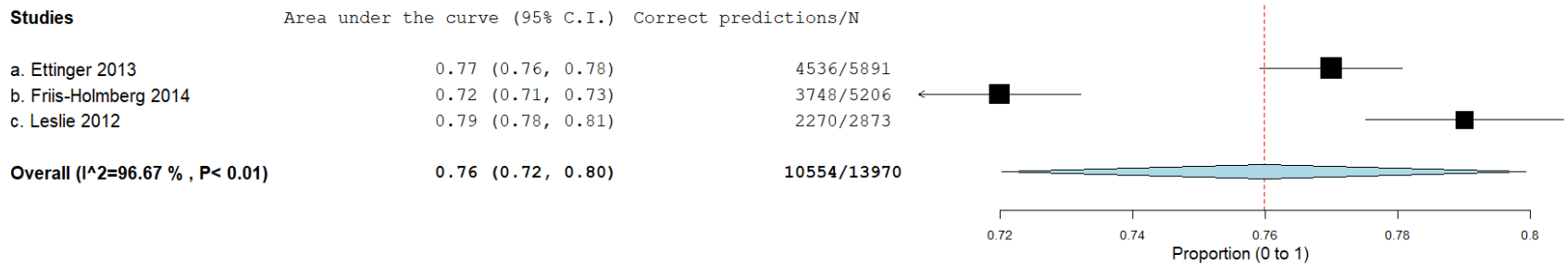
[d] Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. Osteoporos Int. 2012 Jan;23(1):75-85. doi: 10.1007/s00198-011-1747-2. PMID: 21850546.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 13. FRAX with bone mineral density for predicting hip fractures in men (KQ 2a)



[a] 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 Apr;24(4):1185-93. doi: 10.1007/s00198-012-2215-3 [doi]. PMID: 23179575.

[b] 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 Jan-Mar;17(1):7-15. doi: 10.1016/j.jocd.2013.03.014. PMID: 23623379.

[c] Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. Osteoporos Int. 2012 Jan;23(1):75-85. doi: 10.1007/s00198-011-1747-2. PMID: 21850546.

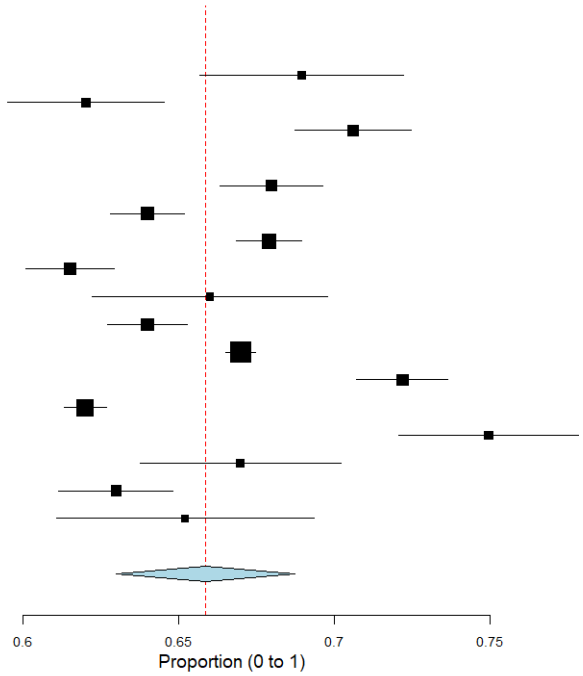
Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 14. FRAX without bone mineral density testing for predicting major osteoporotic fractures in women (KQ 2a)

Studies	Area under the curve (95% C.I.)	Correct predictions/N
a. Azagra 2012	0.69 (0.66, 0.72)	531/770
b. Bolland 2011	0.62 (0.60, 0.65)	882/1422
c. Cheung 2012	0.71 (0.69, 0.72)	1600/2266
d. Crandall 2014	0.56 (0.56, 0.56)	34996/62492
e. Donaldson 2009	0.68 (0.66, 0.70)	2069/3043
f. Ensrud 2009	0.64 (0.63, 0.65)	4001/6252
g. Friis-Holmberg 2014	0.68 (0.67, 0.69)	5128/7552
h. Gonzalez-Macias 2012	0.62 (0.60, 0.63)	2739/4453
i. Henry 2011	0.66 (0.62, 0.70)	396/600
j. Kalveston 2016	0.64 (0.63, 0.65)	3378/5278
k. Leslie 2012	0.67 (0.67, 0.67)	24609/36730
l. Rubin 2013	0.72 (0.71, 0.74)	2609/3614
m. Sambrook 2011	0.62 (0.61, 0.63)	12143/19586
n. Sornay-Rendu 2010	0.75 (0.72, 0.78)	650/867
o. Tamaki 2011	0.67 (0.64, 0.70)	546/815
p. Tremollieres 2010	0.63 (0.61, 0.65)	1670/2651
q. Van Geel 2014	0.65 (0.61, 0.69)	330/506
Overall (I^2=99.16 % , P< 0.01)	**0.66 (0.63, 0.69)**	**98277/158897**



[a] Azagra R, Roca G, Encabo G, et al. FRAX(R) tool, the WHO algorithm to predict osteoporotic fractures: the first analysis of its discriminative and predictive ability in the Spanish FRIDEX cohort. BMC Musculoskelet Disord. 2012;13:204. doi: 10.1186/1471-2474-13-204. PMID: 23088223

[b] Bolland MJ, Siu AT, Mason BH, et al Evaluation of the FRAX and Garvan fracture risk calculators in older women. J Bone Miner Res. 2011 Feb;26(2):420-7. doi: 10.1002/jbmr.215. PMID: 20721930

[c] Cheung EY, Bow CH, Cheung CL, et al Discriminative value of FRAX for fracture prediction in a cohort of Chinese postmenopausal women. Osteoporos Int. 2012 Mar;23(3):871-8. doi: 10.1007/s00198-011-1647-5 [doi]. PMID: 21562875.

[d] Crandall CJ, Larson JC, Watts NB, et al Comparison of fracture risk prediction by the US preventive services task force strategy and two alternative strategies in women 50-64 years old in the women's health initiative. *J Clin Endocrinol Metab*. 2014;99(12):4514-22.

[e] Donaldson MG, Palermo L, Schousboe JT, et al FRAX and risk of vertebral fractures: the fracture intervention trial. *J Bone Miner Res*. 2009 Nov;24(11):1793-9. doi: 10.1359/jbmr.090511 [doi]. PMID: 19419318.

[f] Ensrud KE, Lui LY, Taylor BC, et al A comparison of prediction models for fractures in older women: is more better? *Arch Intern Med*. 2009 Dec 14;169(22):2087-94. doi: 10.1001/archinternmed.2009.404. PMID: 20008691.

[g] 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 Jan-Mar;17(1):7-15. doi: 10.1016/j.jocd.2013.03.014. PMID: 23623379.

[h] Gonzalez-Macias J, Marin F, Vila J, et al Probability of fractures predicted by FRAX(R) and observed incidence in the Spanish ECOSAP Study cohort. *Bone*. 2012 Jan;50(1):373-7. doi: 10.1016/j.bone.2011.11.006. PMID: 22129640.

[i] Henry MJ, Pasco JA, Merriman EN, et al. Fracture risk score and absolute risk of fracture. Radiology. 2011;259(2):495-501. doi: 10.1148/radiol.10101406

© 2018 American Medical Association. All rights reserved.

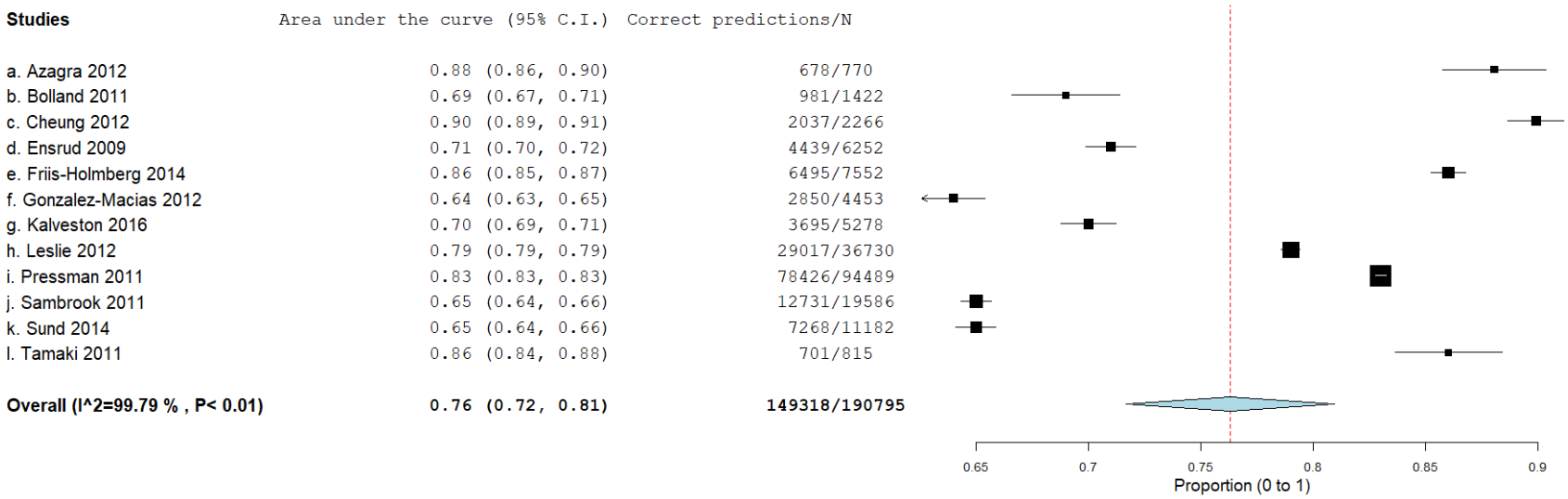
[j] Kalvesten J, Lui LY, Brismar T, Cummings S. Digital X-ray radiogrammetry in the study of osteoporotic fractures: Comparison to dual energy X-ray absorptiometry and FRAX. Bone. 2016:30-5. doi: 10.1016/j.bone.2016.02.011
[k] Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. Osteoporos Int. 2012 Jan;23(1):75-85. doi: 10.1007/s00198-011-1747-2. PMID: 21850546.
[l] Rubin KH, Abrahamsen B, Friis-Holmberg T, et al. Comparison of different screening tools (FRAX(R), OST, ORAI, OSIRIS, SCORE and age alone) to identify women with increased risk of fracture. A population-based prospective study. Bone. 2013;56(1):16-22. doi: 10.1016/j.bone.2013.05.002
[m] Sambrook PN, Flahive J, Hooven FH, et al Predicting fractures in an international cohort using risk factor algorithms without BMD. *J Bone Miner Res*. 2011 Nov;26(11):2770-7. doi: 10.1002/jbmr.503. PMID: 21887705.
[n] Sornay-Rendu E, Munoz F, Delmas PD, et al The FRAX tool in French women: How well does it describe the real incidence of fracture in the OFELY cohort? *J Bone Miner Res*. 2010 Oct;25(10):2101-7. doi: 10.1002/jbmr.106. PMID: 20499352.
[o] Tamaki J, Iki M, Kadowaki E, et al Fracture risk prediction using FRAX(R): a 10-year follow-up survey of the Japanese Population-Based Osteoporosis (JPOS) Cohort Study. *Osteoporos Int*. 2011 Dec;22(12):3037-45. doi: 10.1007/s00198-011-1537-x [doi]. PMID: 21279504.
[p] Tremollieres FA, Pouilles JM, Drewniak N, et al Fracture risk prediction using BMD and clinical risk factors in early postmenopausal women: sensitivity of the WHO FRAX tool. *J Bone Miner Res*. 2010 May;25(5):1002-9. doi: 10.1002/jbmr.12. PMID: 20200927.
[q] van Geel TA, Eisman JA, Geusens PP, van den Bergh JP, Center JR, Dinant GJ. The utility of absolute risk prediction using FRAX(R) and Garvan Fracture Risk Calculator in daily practice. Maturitas. 2014;77(2):174-9. doi: 10.1016/j.maturitas.2013.10.021

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 15. FRAX without bone mineral density testing for predicting hip fractures in women (KQ 2a)



[a] Azagra R, Roca G, Encabo G, et al. FRAX(R) tool, the WHO algorithm to predict osteoporotic fractures: the first analysis of its discriminative and predictive ability in the Spanish FRIDEX cohort. BMC Musculoskelet Disord. 2012;13:204. doi: 10.1186/1471-2474-13-204. PMID: 23088223

[b] Bolland MJ, Siu AT, Mason BH, et al Evaluation of the FRAX and Garvan fracture risk calculators in older women. J Bone Miner Res. 2011 Feb;26(2):420-7. doi: 10.1002/jbmr.215. PMID: 20721930

[c] Cheung EY, Bow CH, Cheung CL, et al Discriminative value of FRAX for fracture prediction in a cohort of Chinese postmenopausal women. Osteoporos Int. 2012 Mar;23(3):871-8. doi: 10.1007/s00198-011-1647-5 [doi]. PMID: 21562875.

[d] Ensrud KE, Lui LY, Taylor BC, et al A comparison of prediction models for fractures in older women: is more better? Arch Intern Med. 2009 Dec 14;169(22):2087-94. doi: 10.1001/archinternmed.2009.404. PMID: 20008691.

[e] 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 Jan-Mar;17(1):7-15. doi: 10.1016/j.jocd.2013.03.014. PMID: 23623379.

[f] Gonzalez-Macias J, Marin F, Vila J, et al Probability of fractures predicted by FRAX(R) and observed incidence in the Spanish ECOSAP Study cohort. Bone. 2012 Jan;50(1):373-7. doi: 10.1016/j.bone.2011.11.006. PMID: 22129640.

[g] Kalvesten J, Lui LY, Brismar T, Cummings S. Digital X-ray radiogrammetry in the study of osteoporotic fractures: Comparison to dual energy X-ray absorptiometry and FRAX. Bone. 2016:30-5. doi: 10.1016/j.bone.2016.02.011

[h] Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. Osteoporos Int. 2012 Jan;23(1):75-85. doi: 10.1007/s00198-011-1747-2. PMID: 21850546.

[i] Pressman AR, Lo JC, Chandra M, et al Methods for assessing fracture risk prediction models: experience with FRAX in a large integrated health care delivery system. J Clin Densitom. 2011 Oct-Dec;14(4):407-15. doi: 10.1016/j.jocd.2011.06.006. PMID: 21958955.

[j] Sambrook PN, Flahive J, Hooven FH, et al Predicting fractures in an international cohort using risk factor algorithms without BMD. J Bone Miner Res. 2011 Nov;26(11):2770-7. doi: 10.1002/jbmr.503. PMID: 21887705.

[k] Sund R, Honkanen R, Johansson H, et al Evaluation of the FRAX model for hip fracture predictions in the population-based Kuopio Osteoporosis Risk Factor and Prevention Study (OSTPRE). Calcif Tissue Int. 2014 Jul;95(1):39-45. doi: 10.1007/s00223-014-9860-9 [doi]. PMID: 24792689.

© 2018 American Medical Association. All rights reserved.

[1] Tamaki J, Iki M, Kadowaki E, et al Fracture risk prediction using FRAX(R): a 10-year follow-up survey of the Japanese Population-Based Osteoporosis (JPOS) Cohort Study. Osteoporos Int. 2011 Dec;22(12):3037-45. doi: 10.1007/s00198-011-1537-x [doi]. PMID: 21279504.

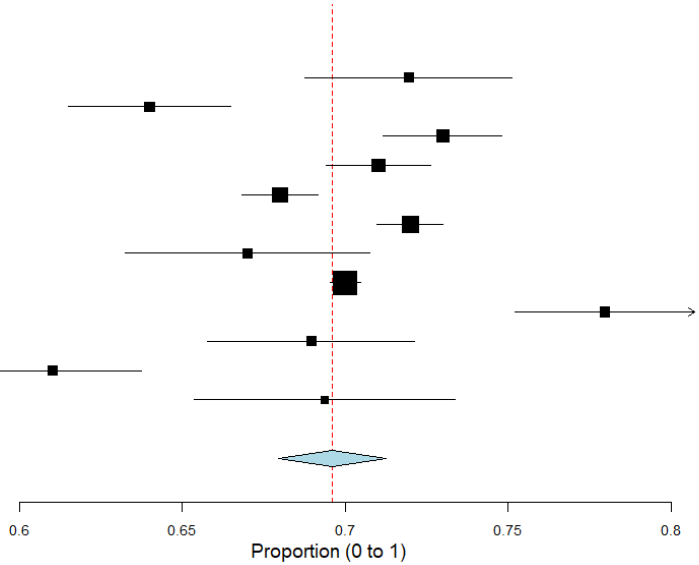
Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate

© 2018 American Medical Association. All rights reserved.

eFigure 16. FRAX with bone mineral density testing for predicting major osteoporotic fractures in women (KQ 2a)

Studies	Area under the curve (95% C.I.)	Correct predictions/N
a. Azagra 2012	0.72 (0.69, 0.75)	554/770
b. Bolland 2011	0.64 (0.61, 0.66)	910/1422
c. Cheung 2012	0.73 (0.71, 0.75)	1654/2266
d. Donaldson 2009	0.71 (0.69, 0.73)	2161/3043
e. Ensrud 2009	0.68 (0.67, 0.69)	4251/6252
f. Friis-Holmberg 2014	0.72 (0.71, 0.73)	5437/7552
g. Henry 2011	0.67 (0.63, 0.71)	402/600
h. Leslie 2012	0.70 (0.70, 0.70)	25711/36730
i. Sornay-Rendu 2010	0.78 (0.75, 0.81)	676/867
j. Tamaki 2011	0.69 (0.66, 0.72)	562/815
k. Tebe-Cordomi 2013	0.61 (0.58, 0.64)	751/1231
l. Van Geel 2014	0.69 (0.65, 0.73)	351/506
Overall (I^2=92.07 % , P< 0.01)	**0.70 (0.68, 0.71)**	**43420/62054**



[a] Azagra R, Roca G, Encabo G, et al. FRAX(R) tool, the WHO algorithm to predict osteoporotic fractures: the first analysis of its discriminative and predictive ability in the Spanish FRIDEX cohort. BMC Musculoskelet Disord. 2012;13:204. doi: 10.1186/1471-2474-13-204. PMID: 23088223.

[b] Bolland MJ, Siu AT, Mason BH, et al Evaluation of the FRAX and Garvan fracture risk calculators in older women. J Bone Miner Res. 2011 Feb;26(2):420-7. doi: 10.1002/jbmr.215. PMID: 20721930

[c] Cheung EY, Bow CH, Cheung CL, et al Discriminative value of FRAX for fracture prediction in a cohort of Chinese postmenopausal women. Osteoporos Int. 2012 Mar;23(3):871-8. doi: 10.1007/s00198-011-1647-5 [doi]. PMID: 21562875.

[d] Donaldson MG, Palermo L, Schousboe JT, et al FRAX and risk of vertebral fractures: the fracture intervention trial. *J Bone Miner Res*. 2009 Nov;24(11):1793-9. doi: 10.1359/jbmr.090511 [doi]. PMID: 19419318.

[e] Ensrud KE, Lui LY, Taylor BC, et al A comparison of prediction models for fractures in older women: is more better? *Arch Intern Med*. 2009 Dec 14;169(22):2087-94. doi: 10.1001/archinternmed.2009.404. PMID: 20008691.

[f] 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 Jan-Mar;17(1):7-15. doi: 10.1016/j.jocd.2013.03.014. PMID: 23623379.

[g] Henry MJ, Pasco JA, Merriman EN, et al. Fracture risk score and absolute risk of fracture. Radiology. 2011;259(2):495-501. doi: 10.1148/radiol.10101406

[h] Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. Osteoporos Int. 2012 Jan;23(1):75-85. doi: 10.1007/s00198-011-1747-2. PMID: 21850546.

[i] Sornay-Rendu E, Munoz F, Delmas PD, et al The FRAX tool in French women: How well does it describe the real incidence of fracture in the OFELY cohort? *J Bone Miner Res*. 2010 Oct;25(10):2101-7. doi: 10.1002/jbmr.106. PMID: 20499352.

[j] Tamaki J, Iki M, Kadowaki E, et al Fracture risk prediction using FRAX(R): a 10-year follow-up survey of the Japanese Population-Based Osteoporosis (JPOS) Cohort Study. *Osteoporos Int*. 2011 Dec;22(12):3037-45. doi: 10.1007/s00198-011-1537-x [doi]. PMID: 21279504.

[k] Tebe Cordomi C, Del Rio LM, Di Gregorio S, et al Validation of the FRAX predictive model for major osteoporotic fracture in a historical cohort of Spanish women. *J Clin Densitom*. 2013 Apr-Jun;16(2):231-7. doi: 10.1016/j.jocd.2012.05.007. PMID: 22748778.

© 2018 American Medical Association. All rights reserved.

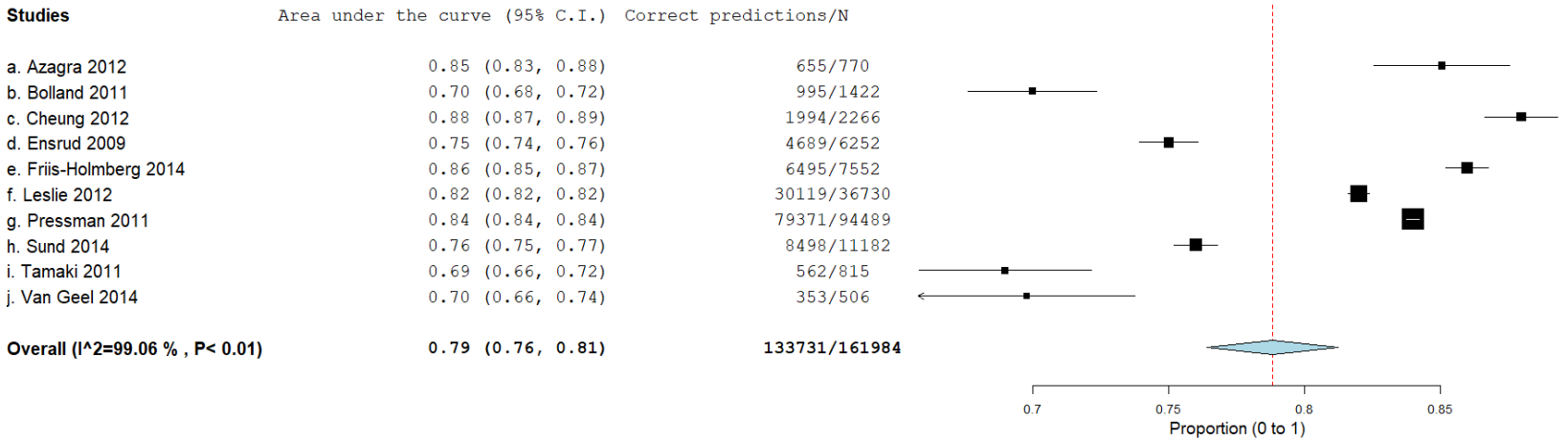
[l] van Geel TA, Eisman JA, Geusens PP, van den Bergh JP, Center JR, Dinant GJ. The utility of absolute risk prediction using FRAX(R) and Garvan Fracture Risk Calculator in daily practice. Maturitas. 2014;77(2):174-9. doi: 10.1016/j.maturitas.2013.10.021

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate

© 2018 American Medical Association. All rights reserved.

eFigure 17. FRAX with bone mineral density testing for predicting hip fractures in women (KQ 2a)



[a] Azagra R, Roca G, Encabo G, et al. FRAX(R) tool, the WHO algorithm to predict osteoporotic fractures: the first analysis of its discriminative and predictive ability in the Spanish FRIDEX cohort. BMC Musculoskelet Disord. 2012;13:204. doi: 10.1186/1471-2474-13-204. PMID: 23088223.

[b] Bolland MJ, Siu AT, Mason BH, et al Evaluation of the FRAX and Garvan fracture risk calculators in older women. J Bone Miner Res. 2011 Feb;26(2):420-7. doi: 10.1002/jbmr.215. PMID: 20721930

[c] Cheung EY, Bow CH, Cheung CL, et al Discriminative value of FRAX for fracture prediction in a cohort of Chinese postmenopausal women. Osteoporos Int. 2012 Mar;23(3):871-8. doi: 10.1007/s00198-011-1647-5 [doi]. PMID: 21562875.

[d] Ensrud KE, Lui LY, Taylor BC, et al A comparison of prediction models for fractures in older women: is more better? *Arch Intern Med*. 2009 Dec 14;169(22):2087-94. doi: 10.1001/archinternmed.2009.404. PMID: 20008691.

[e] 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 Jan-Mar;17(1):7-15. doi: 10.1016/j.jocd.2013.03.014. PMID: 23623379.

[f] Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. Osteoporos Int. 2012 Jan;23(1):75-85. doi: 10.1007/s00198-011-1747-2. PMID: 21850546.

[g] Pressman AR, Lo JC, Chandra M, et al Methods for assessing fracture risk prediction models: experience with FRAX in a large integrated health care delivery system. *J Clin Densitom*. 2011 Oct-Dec;14(4):407-15. doi: 10.1016/j.jocd.2011.06.006. PMID: 21958955.

[h] Sund R, Honkanen R, Johansson H, et al Evaluation of the FRAX model for hip fracture predictions in the population-based Kuopio Osteoporosis Risk Factor and Prevention Study (OSTPRE). *Calcif Tissue Int*. 2014 Jul;95(1):39-45. doi: 10.1007/s00223-014-9860-9 [doi]. PMID: 24792689.

[i] Tamaki J, Iki M, Kadowaki E, et al Fracture risk prediction using FRAX(R): a 10-year follow-up survey of the Japanese Population-Based Osteoporosis (JPOS) Cohort Study. *Osteoporos Int*. 2011 Dec;22(12):3037-45. doi: 10.1007/s00198-011-1537-x [doi]. PMID: 21279504.

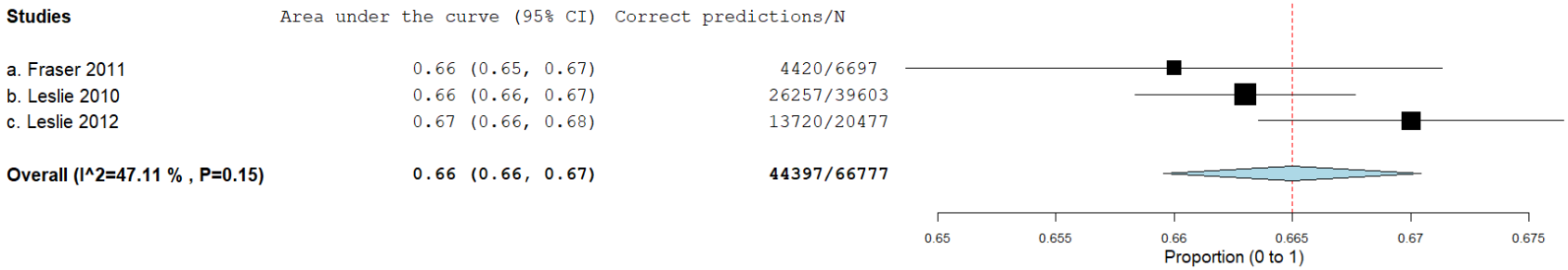
[j] van Geel TA, Eisman JA, Geusens PP, van den Bergh JP, Center JR, Dinant GJ. The utility of absolute risk prediction using FRAX(R) and Garvan Fracture Risk Calculator in daily practice. Maturitas. 2014;77(2):174-9. doi: 10.1016/j.maturitas.2013.10.021

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 18. FRAX without bone mineral density testing for predicting major osteoporotic fractures in both sexes (KQ 2a)



[a] Fraser LA, Langsetmo L, Berger C, et al Fracture prediction and calibration of a Canadian FRAX(R) tool: a population-based report from CaMos. *Osteoporos Int*. 2011 Mar;22(3):829-37. doi: 10.1007/s00198-010-1465-1 [doi]. PMID: 21161508

[b] Leslie WD, Lix LM, Johansson H, et al Independent clinical validation of a Canadian FRAX tool: fracture prediction and model calibration. *J Bone Miner Res*. 2010 Nov;25(11):2350-8. doi: 10.1002/jbmr.123 [doi]. PMID: 20499367.

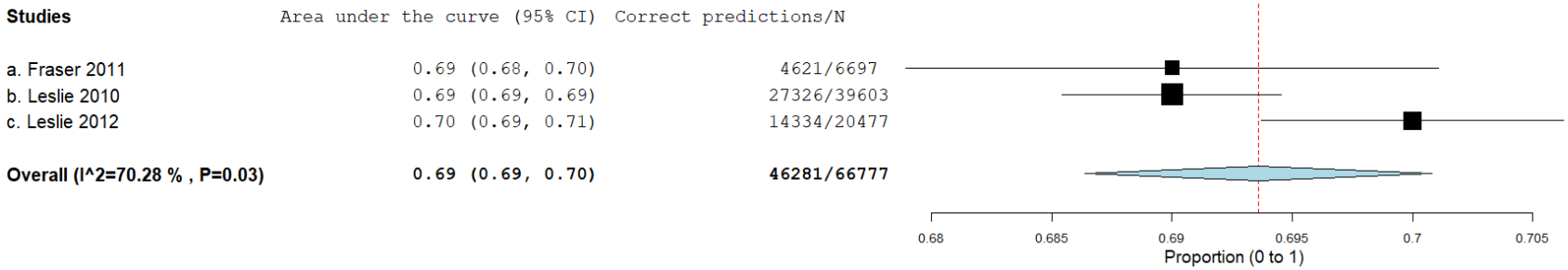
[c] Leslie WD, Lix LM, Johansson H, Oden A, McCloskey E, Kanis JA. A comparative study of using non-hip bone density inputs with FRAX(R). Osteoporos Int. 2012;23(3):853-60. doi: 10.1007/s00198-011-1814-8 [doi]

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate

© 2018 American Medical Association. All rights reserved.

eFigure 19. FRAX with bone mineral density testing for predicting major osteoporotic fractures in both sexes (KQ 2a)



[a] Fraser LA, Langsetmo L, Berger C, et al Fracture prediction and calibration of a Canadian FRAX(R) tool: a population-based report from CaMos. Osteoporos Int. 2011 Mar;22(3):829-37. doi: 10.1007/s00198-010-1465-1 [doi]. PMID: 21161508

[b] Leslie WD, Lix LM, Johansson H, et al Independent clinical validation of a Canadian FRAX tool: fracture prediction and model calibration. J Bone Miner Res. 2010 Nov;25(11):2350-8. doi: 10.1002/jbmr.123 [doi]. PMID: 20499367.

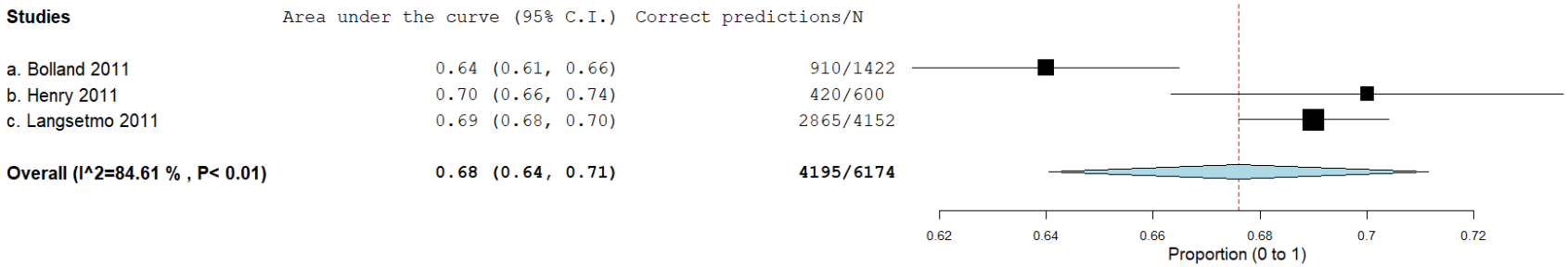
[c] Leslie WD, Lix LM, Johansson H, Oden A, McCloskey E, Kanis JA. A comparative study of using non-hip bone density inputs with FRAX(R). Osteoporos Int. 2012;23(3):853-60. doi: 10.1007/s00198-011-1814-8 [doi]

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate

© 2018 American Medical Association. All rights reserved.

eFigure 20. Garvan Fracture Risk Calculator with bone mineral density testing for predicting major osteoporotic fractures in women (KQ 2a)



[a] Bolland MJ, Siu AT, Mason BH, et al Evaluation of the FRAX and Garvan fracture risk calculators in older women. *J Bone Miner Res*. 2011 Feb;26(2):420-7. doi: 10.1002/jbmr.215. PMID: 20721930.

[b] Henry MJ, Pasco JA, Merriman EN, et al. Fracture risk score and absolute risk of fracture. Radiology. 2011;259(2):495-501. doi: 10.1148/radiol.10101406

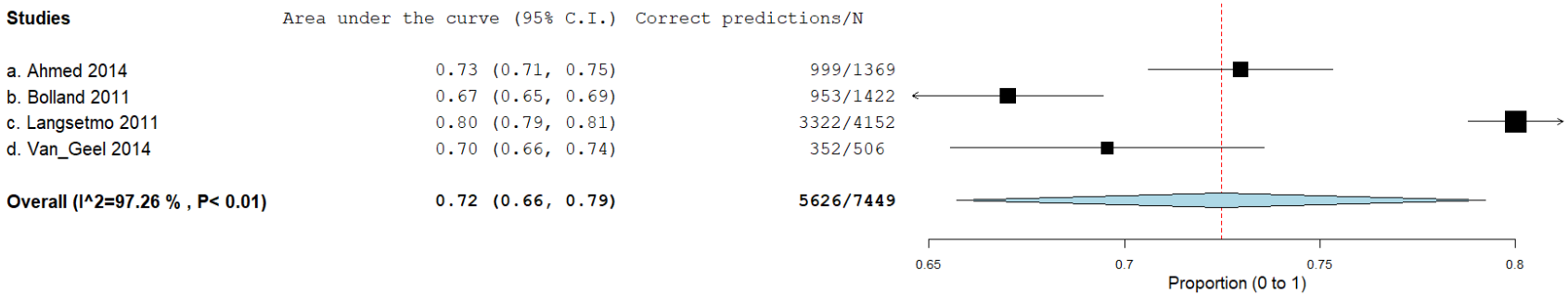
[c] Langsetmo L, Nguyen TV, Nguyen ND, et al Independent external validation of nomograms for predicting risk of low-trauma fracture and hip fracture. *CMAJ*. 2011 Feb 8;183(2):E107-14. doi: 10.1503/cmaj.100458. PMID: 21173069.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate

© 2018 American Medical Association. All rights reserved.

eFigure 21. Garvan Fracture Risk Calculator with bone mineral density testing for predicting hip fractures in women (KQ 2a)



[a] Bolland MJ, Siu AT, Mason BH, et al Evaluation of the FRAX and Garvan fracture risk calculators in older women. *J Bone Miner Res*. 2011 Feb;26(2):420-7. doi: 10.1002/jbmr.215. PMID: 20721930.

[b] Henry MJ, Pasco JA, Merriman EN, et al. Fracture risk score and absolute risk of fracture. Radiology. 2011;259(2):495-501. doi: 10.1148/radiol.10101406

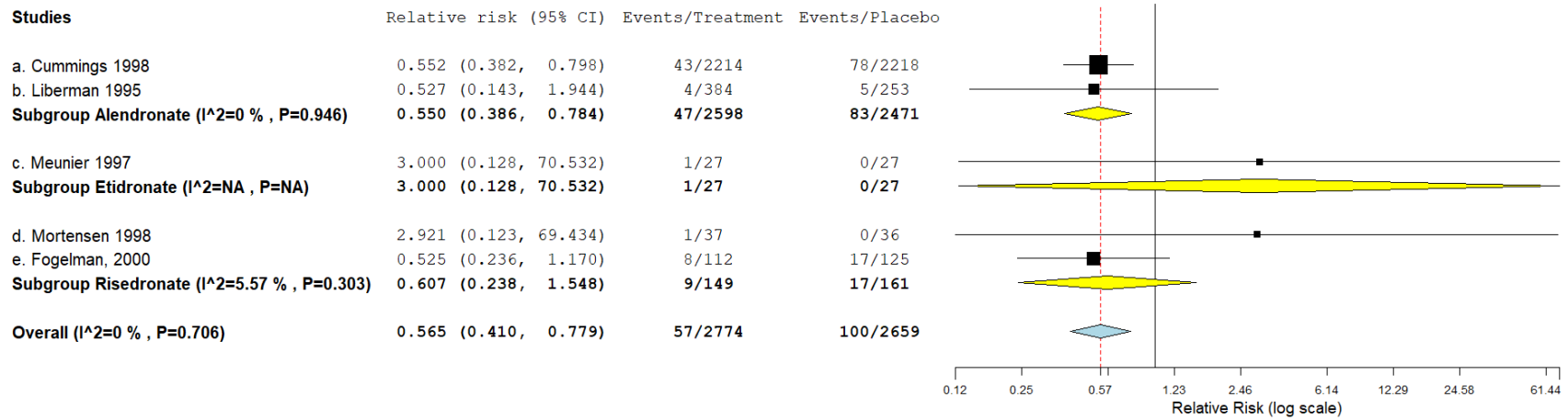
[c] Langsetmo L, Nguyen TV, Nguyen ND, et al Independent external validation of nomograms for predicting risk of low-trauma fracture and hip fracture. *CMAJ*. 2011 Feb 8;183(2):E107-14. doi: 10.1503/cmaj.100458. PMID: 21173069.

Abbreviations: AUC=confidence interval; N=number

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of accuracy; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate

© 2018 American Medical Association. All rights reserved.

eFigure 22. Vertebral fracture outcomes for bisphosphonates (KQ 4a)



[a] Cummings SR, Black DM, Thompson DE, et al Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA*. 1998 Dec 23-30;280(24):2077-82. PMID: 9875874.

[b] Liberman UA, Weiss SR, Broll J, et al Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. The Alendronate Phase III Osteoporosis Treatment Study Group. *N Engl J Med*. 1995 Nov 30;333(22):1437-43. doi: 10.1056/nejm199511303332201. PMID: 7477143.

[c] Meunier PJ, Confavreux E, Tupinon I, et al Prevention of early postmenopausal bone loss with cyclical etidronate therapy (a double-blind, placebo-controlled study and 1-year follow-up). *J Clin Endocrinol Metab*. 1997 Sep;82(9):2784-91. doi: 10.1210/jcem.82.9.4073. PMID: 9284696.

[d] Mortensen L, Charles P, Bekker PJ, et al Risedronate increases bone mass in an early postmenopausal population: two years of treatment plus one year of follow-up. *J Clin Endocrinol Metab*. 1998 Feb;83(2):396-402. doi: 10.1210/jcem.83.2.4586. PMID: 9467547.

[e] Fogelman I, Ribot C, Smith R, et al Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. *J Clin Endocrinol Metab*. 2000 May;85(5):1895-900. doi: 10.1210/jcem.85.5.6603. PMID: 10843171.

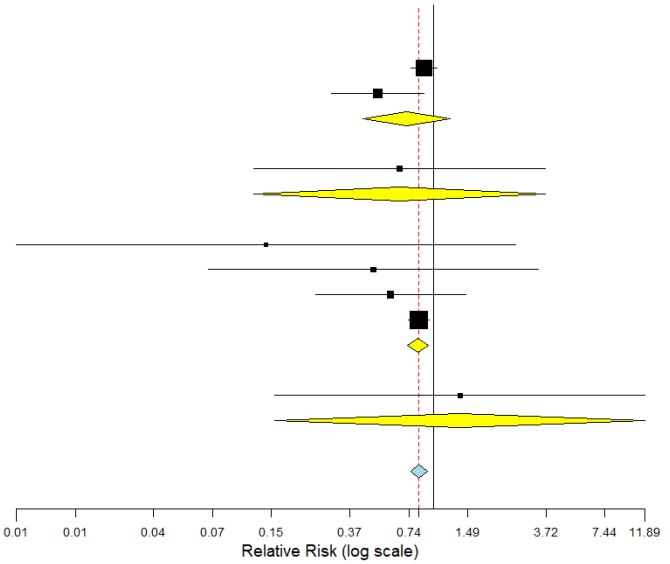
Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 23. Nonvertebral fracture outcomes for bisphosphonates (KQ 4a)

Studies	Relative risk (95% CI)	Events/Treatment	Events/Placebo
a. Cummings 1998	0.889 (0.761, 1.039)	261/2214	294/2218
b. Pols 1999	0.518 (0.300, 0.894)	19/950	37/958
Subgroup Alendronate (I^2=71.32 % , P=0.062)	**0.725 (0.433, 1.212)**	**280/3164**	**331/3176**
c. Meunier 1997	0.667 (0.121, 3.678)	2/27	3/27
Subgroup Etidronate (I^2=NA , P=NA)	**0.667 (0.121, 3.678)**	**2/27**	**3/27**
d. Mortensen 1998	0.139 (0.007, 2.601)	0/37	3/36
e. Valimaiki 2007	0.491 (0.071, 3.397)	2/114	2/56
f. Fogelman 2000	0.601 (0.249, 1.453)	7/112	13/125
g. McClung 2001	0.839 (0.740, 0.950)	582/6197	351/3134
Subgroup Risedronate (I^2=0 % , P=0.522)	**0.829 (0.732, 0.938)**	**591/6460**	**369/3351**
h. Reid 2002	1.356 (0.155, 11.895)	4/174	1/59
Subgroup Zoledronic Acid (I^2=NA , P=NA)	**1.356 (0.155, 11.895)**	**4/174**	**1/59**
Overall (I^2=0 % , P=0.530)	**0.839 (0.763, 0.923)**	**877/9825**	**704/6613**



[a] Cummings SR, Black DM, Thompson DE, et al Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA*. 1998 Dec 23-30;280(24):2077-82. PMID: 9875874.

[b] Pols HA, Felsenberg D, Hanley DA, et al Multinational, placebo-controlled, randomized trial of the effects of alendronate on bone density and fracture risk in postmenopausal women with low bone mass: results of the FOSIT study. Fosamax International Trial Study Group. Osteoporos Int. 1999;9(5):461-8. PMID: 10550467.

[c] Meunier PJ, Confavreux E, Tupinon I, et al Prevention of early postmenopausal bone loss with cyclical etidronate therapy (a double-blind, placebo-controlled study and 1-year follow-up). *J Clin Endocrinol Metab*. 1997 Sep;82(9):2784-91. doi: 10.1210/jcem.82.9.4073. PMID: 9284696.

[d] Mortensen L, Charles P, Bekker PJ, et al Risedronate increases bone mass in an early postmenopausal population: two years of treatment plus one year of follow-up. *J Clin Endocrinol Metab*. 1998 Feb;83(2):396-402. doi: 10.1210/jcem.83.2.4586. PMID: 9467547.

[e] Valimaki MJ, Farrerons-Minguella J, Halse J, et al Effects of risedronate 5 mg/d on bone mineral density and bone turnover markers in late-postmenopausal women with osteopenia: a multinational, 24-month, randomized, double-blind, placebo-controlled, parallel-group, phase III trial. Clin Ther. 2007 Sep;29(9):1937-49. doi: 10.1016/j.clinthera.2007.09.017. PMID: 18035193.

[f] Fogelman I, Ribot C, Smith R, et al Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. *J Clin Endocrinol Metab*. 2000 May;85(5):1895-900. doi: 10.1210/jcem.85.5.6603. PMID: 10843171.

[g] McClung MR, Geusens P, Miller PD, et al Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. N Engl J Med. 2001 Feb 1;344(5):333-40. doi: 10.1056/nejm200102013440503. PMID: 11172164.

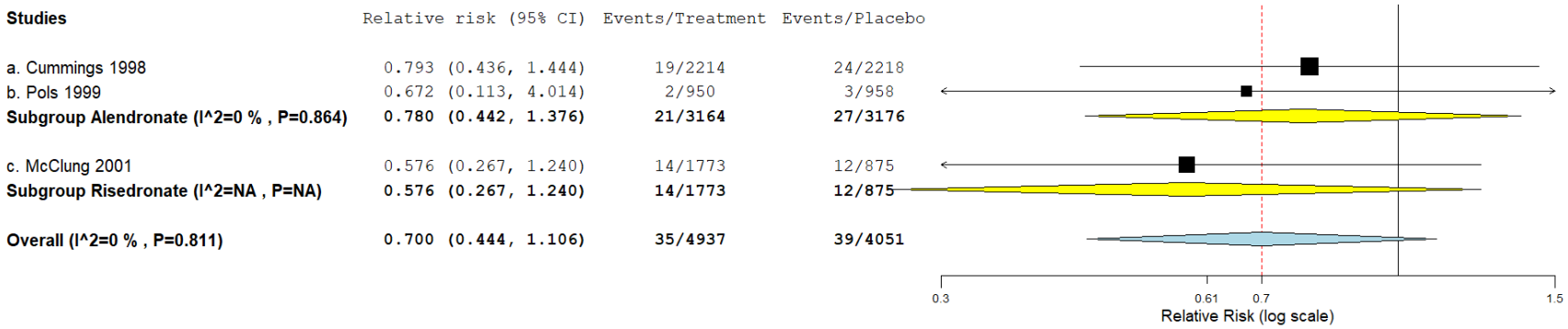
[h] Reid IR, Brown JP, Burckhardt P, et al Intravenous zoledronic acid in postmenopausal women with low bone mineral density. N Engl J Med. 2002 Feb 28;346(9):653-61. doi: 10.1056/NEJMoa011807. PMID: 11870242.

Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 24. Hip fracture outcomes for bisphosphonates (KQ 4a)



[a] Cummings SR, Black DM, Thompson DE, et al Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA*. 1998 Dec 23-30;280(24):2077-82. PMID: 9875874.

[b] Pols HA, Felsenberg D, Hanley DA, et al Multinational, placebo-controlled, randomized trial of the effects of alendronate on bone density and fracture risk in postmenopausal women with low bone mass: results of the FOSIT study. Fosamax International Trial Study Group. Osteoporos Int. 1999;9(5):461-8. PMID: 10550467.

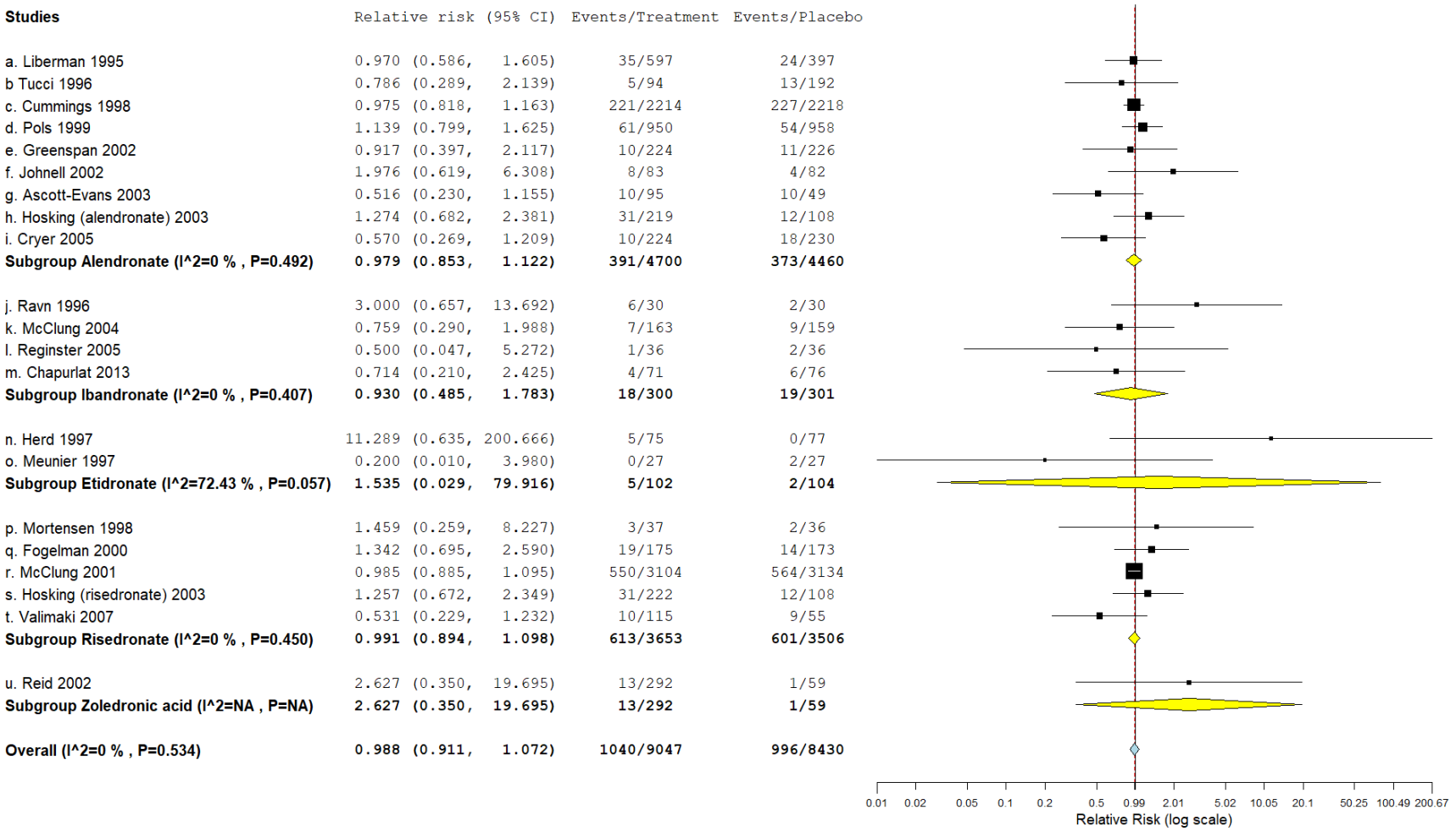
[c] McClung MR, Geusens P, Miller PD, et al Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. N Engl J Med. 2001 Feb 1;344(5):333-40. doi: 10.1056/nejm200102013440503. PMID: 11172164.

Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 25. Discontinuation due to adverse events for bisphosphonates versus placebo (KQ5)



[a] Liberman UA, Weiss SR, Broll J, et al Effect of oral alendronate on bone mineral density and the incidence of fractures in postmenopausal osteoporosis. The Alendronate Phase III Osteoporosis Treatment Study Group. *N Engl J Med*. 1995 Nov 30;333(22):1437-43. doi: 10.1056/nejm199511303332201. PMID: 7477143.

[b] Tucci JR, Tonino RP, Emkey RD, et al Effect of three years of oral alendronate treatment in postmenopausal women with osteoporosis. *Am J Med*. 1996 Nov;101(5):488-501. PMID: 8948272.

[c] Cummings SR, Black DM, Thompson DE, et al Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA*. 1998 Dec 23-30;280(24):2077-82. PMID: 9875874.

[d] Pols HA, Felsenberg D, Hanley DA, et al Multinational, placebo-controlled, randomized trial of the effects of alendronate on bone density and fracture risk in postmenopausal women with low bone mass: results of the FOSIT study. Fosamax International Trial Study Group. *Osteoporos Int*. 1999;9(5):461-8. PMID: 10550467.

© 2018 American Medical Association. All rights reserved.

[e] Greenspan S, Field-Munves E, Tonino R, et al Tolerability of once-weekly alendronate in patients with osteoporosis: a randomized, double-blind, placebo-controlled study. *Mayo Clin Proc*. 2002 Oct;77(10):1044-52. doi: 10.4065/77.10.1044. PMID: 12374248.
[f] Johnell O, Scheele WH, Lu Y, et al Additive effects of raloxifene and alendronate on bone density and biochemical markers of bone remodeling in postmenopausal women with osteoporosis. *J Clin Endocrinol Metab*. 2002 Mar;87(3):985-92. doi: 10.1210/jcem.87.3.8325. PMID: 11889149.
[g] Ascott-Evans BH, Guanabens N, Kivinen S, et al Alendronate prevents loss of bone density associated with discontinuation of hormone replacement therapy: a randomized controlled trial. *Arch Intern Med*. 2003 Apr 14;163(7):789-94. doi: 10.1001/archinte.163.7.789. PMID: 12695269.
[h] Hosking D, Adami S, Felsenberg D, et al Comparison of change in bone resorption and bone mineral density with once-weekly alendronate and daily risedronate: a randomised, placebo-controlled study. *Curr Med Res Opin*. 2003;19(5):383-94. doi: 10.1185/030079903125002009. PMID: 13678475.
[i] Cryer B, Binkley N, Simonelli C, et al A randomized, placebo-controlled, 6-month study of once-weekly alendronate oral solution for postmenopausal osteoporosis. Am J Geriatr Pharmacother. 2005 Sep;3(3):127-36. PMID: 16257815.
[j] Ravn P, Clemmesen B, Riis BJ, et al The effect on bone mass and bone markers of different doses of ibandronate: a new bisphosphonate for prevention and treatment of postmenopausal osteoporosis: a 1-year, randomized, double-blind, placebo-controlled dose-finding study. *Bone*. 1996 Nov;19(5):527-33. PMID: 8922653.
[k] McClung MR, Wasnich RD, Recker R, et al Oral daily ibandronate prevents bone loss in early postmenopausal women without osteoporosis. *J Bone Miner Res*. 2004 Jan;19(1):11-8. doi: 10.1359/jbmr.0301202. PMID: 14753731.
[l] Reginster JY, Wilson KM, Dumont E, et al Monthly oral ibandronate is well tolerated and efficacious in postmenopausal women: results from the monthly oral pilot study. *J Clin Endocrinol Metab*. 2005 Sep;90(9):5018-24. doi: 10.1210/jc.2004-1750. PMID: 15972582.
[m] Chapurlat RD, Laroche M, Thomas T, et al Effect of oral monthly ibandronate on bone microarchitecture in women with osteopenia-a randomized placebo-controlled trial. *Osteoporos Int*. 2013 Jan;24(1):311-20. doi: 10.1007/s00198-012-1947-4 [doi]. PMID: 22402673.
[n] Herd RJ, Balena R, Blake GM, et al The prevention of early postmenopausal bone loss by cyclical etidronate therapy: a 2-year, double-blind, placebo-controlled study. *Am J Med*. 1997 Aug;103(2):92-9. PMID: 9274891.
[o] Meunier PJ, Confavreux E, Tupinon I, et al Prevention of early postmenopausal bone loss with cyclical etidronate therapy (a double-blind, placebo-controlled study and 1-year follow-up). *J Clin Endocrinol Metab*. 1997 Sep;82(9):2784-91. doi: 10.1210/jcem.82.9.4073. PMID: 9284696.
[p] Mortensen L, Charles P, Bekker PJ, et al Risedronate increases bone mass in an early postmenopausal population: two years of treatment plus one year of follow-up. *J Clin Endocrinol Metab*. 1998 Feb;83(2):396-402. doi: 10.1210/jcem.83.2.4586. PMID: 9467547.
[q] Fogelman I, Ribot C, Smith R, et al Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. *J Clin Endocrinol Metab*. 2000 May;85(5):1895-900. doi: 10.1210/jcem.85.5.6603. PMID: 10843171.
[r] McClung MR, Geusens P, Miller PD, et al Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. *N Engl J Med*. 2001 Feb 1;344(5):333-40. doi: 10.1056/nejm200102013440503. PMID: 11172164.
[s] Hosking D, Adami S, Felsenberg D, et al Comparison of change in bone resorption and bone mineral density with once-weekly alendronate and daily risedronate: a randomised, placebo-controlled study. *Curr Med Res Opin*. 2003;19(5):383-94. doi: 10.1185/030079903125002009. PMID: 13678475.
[t] Valimaki MJ, Farrerons-Minguella J, Halse J, et al Effects of risedronate 5 mg/d on bone mineral density and bone turnover markers in late-postmenopausal women with osteopenia: a multinational, 24-month, randomized, double-blind, placebo-controlled, parallel-group, phase III trial. *Clin Ther*. 2007 Sep;29(9):1937-49. doi: 10.1016/j.clinthera.2007.09.017. PMID: 18035193.
[u] Reid IR, Brown JP, Burckhardt P, et al Intravenous zoledronic acid in postmenopausal women with low bone mineral density. *N Engl J Med*. 2002 Feb 28;346(9):653-61. doi: 10.1056/NEJMoa011807. PMID: 11870242.

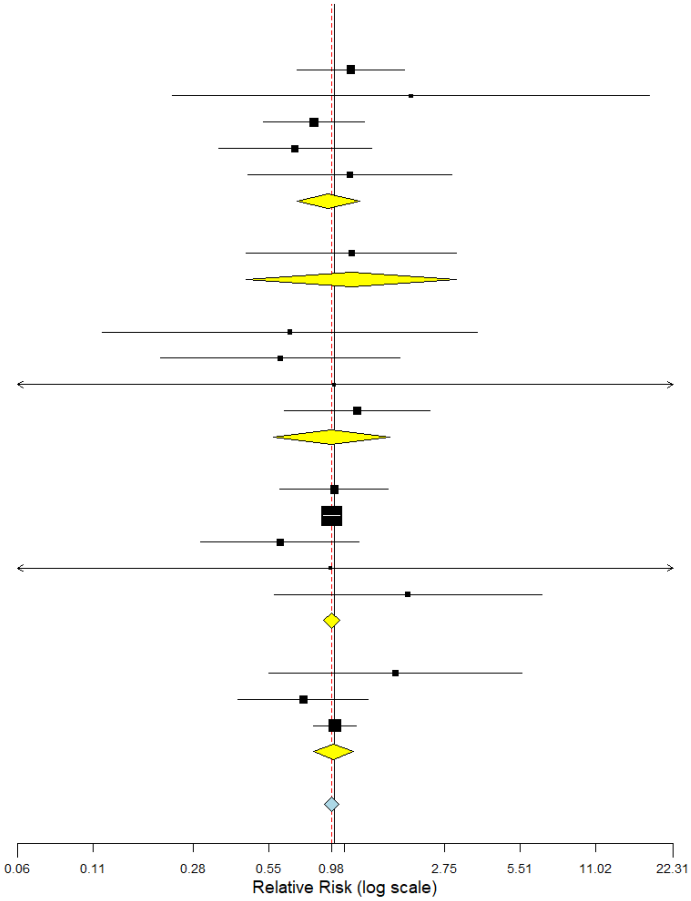
Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 26. Serious adverse events for bisphosphonates versus placebo (KQ5)

Studies	Relative risk (95% CI)	Events/Treatment	Events/Placebo
a. Tucci 1996	1.167 (0.714, 1.907)	20/94	35/192
b. Adachi 2001	2.021 (0.228, 17.916)	4/291	1/147
c. Greenspan 2002	0.831 (0.522, 1.323)	28/224	34/226
d. Hosking (alendronate) 2003	0.699 (0.346, 1.410)	17/219	12/108
e. Cryer 2005	1.155 (0.454, 2.941)	9/224	8/230
Subgroup Alendronate (I^2=0 % , P=0.670)	**0.949 (0.713, 1.265)**	**78/1052**	**90/903**
f. Herd 1997	1.173 (0.448, 3.074)	8/75	7/77
Subgroup Etidronate (I^2=NA , P=NA)	**1.173 (0.448, 3.074)**	**8/75**	**7/77**
g. Ravn 1996	0.667 (0.120, 3.709)	2/30	3/30
h. McClung 2004	0.610 (0.204, 1.824)	5/163	8/159
i. Reginster 2005	1.000 (0.020, 49.075)	0/36	0/36
j. Chapurlat 2013	1.235 (0.633, 2.410)	15/71	13/76
Subgroup Ibandronate (I^2=0 % , P=0.712)	**0.978 (0.572, 1.672)**	**22/300**	**24/301**
k. Fogelman 2000	1.002 (0.610, 1.646)	26/173	27/180
l. McClung 2001	0.979 (0.908, 1.054)	943/3104	973/3134
m. Hosking (risedronate) 2003	0.608 (0.295, 1.254)	15/222	12/108
n. Shiraki 2003	0.963 (0.019, 47.639)	0/53	0/51
o. Valimaki 2007	1.965 (0.578, 6.683)	12/114	3/56
Subgroup Risedronate (I^2=0 % , P=0.572)	**0.977 (0.908, 1.051)**	**996/3666**	**1015/3529**
p. Reid 2002	1.751 (0.548, 5.597)	26/292	3/59
q. McClung 2009	0.754 (0.416, 1.368)	17/198	23/202
r. Boonen 2012	1.005 (0.828, 1.221)	149/588	154/611
Subgroup Zoledronic acid (I^2=0 % , P=0.416)	**0.992 (0.826, 1.191)**	**192/1078**	**180/872**
Overall (I^2=0 % , P=0.952)	**0.978 (0.916, 1.044)**	**1296/6171**	**1316/5682**



[a] Tucci JR, Tonino RP, Emkey RD, et al Effect of three years of oral alendronate treatment in postmenopausal women with osteoporosis. *Am J Med*. 1996 Nov;101(5):488-501. PMID: 8948272.

[b] Adachi JD, Saag KG, Delmas PD, et al Two-year effects of alendronate on bone mineral density and vertebral fracture in patients receiving glucocorticoids: a randomized, double-blind, placebo-controlled extension trial. Arthritis Rheum. 2001 Jan;44(1):202-11. doi: 10.1002/1529-0131(200101)44:1<202::aid-anr27>3.0.co;2-w. PMID: 11212161.

[c] Greenspan S, Field-Munves E, Tonino R, et al Tolerability of once-weekly alendronate in patients with osteoporosis: a randomized, double-blind, placebo-controlled study. *Mayo Clin Proc*. 2002 Oct;77(10):1044-52. doi: 10.4065/77.10.1044. PMID: 12374248.

[d] Hosking D, Adami S, Felsenberg D, et al Comparison of change in bone resorption and bone mineral density with once-weekly alendronate and daily risedronate: a randomised, placebo-controlled study. *Curr Med Res Opin*. 2003;19(5):383-94. doi: 10.1185/030079903125002009. PMID: 13678475.

© 2018 American Medical Association. All rights reserved.

[e] Cryer B, Binkley N, Simonelli C, et al A randomized, placebo-controlled, 6-month study of once-weekly alendronate oral solution for postmenopausal osteoporosis. Am J Geriatr Pharmacother. 2005 Sep;3(3):127-36. PMID: 16257815.
[f] Herd RJ, Balena R, Blake GM, et al The prevention of early postmenopausal bone loss by cyclical etidronate therapy: a 2-year, double-blind, placebo-controlled study. *Am J Med*. 1997 Aug;103(2):92-9. PMID: 9274891.
[g] Ravn P, Clemmesen B, Riis BJ, et al The effect on bone mass and bone markers of different doses of ibandronate: a new bisphosphonate for prevention and treatment of postmenopausal osteoporosis: a 1-year, randomized, double-blind, placebo-controlled dose-finding study. *Bone*. 1996 Nov;19(5):527-33. PMID: 8922653.
[h] McClung MR, Wasnich RD, Recker R, et al Oral daily ibandronate prevents bone loss in early postmenopausal women without osteoporosis. *J Bone Miner Res*. 2004 Jan;19(1):11-8. doi: 10.1359/jbmr.0301202. PMID: 14753731.
[i] Reginster JY, Wilson KM, Dumont E, et al Monthly oral ibandronate is well tolerated and efficacious in postmenopausal women: results from the monthly oral pilot study. *J Clin Endocrinol Metab*. 2005 Sep;90(9):5018-24. doi: 10.1210/jc.2004-1750. PMID: 15972582.
[j] Chapurlat RD, Laroche M, Thomas T, et al Effect of oral monthly ibandronate on bone microarchitecture in women with osteopenia-a randomized placebo-controlled trial. *Osteoporos Int*. 2013 Jan;24(1):311-20. doi: 10.1007/s00198-012-1947-4 [doi]. PMID: 22402673.
[k] Fogelman I, Ribot C, Smith R, et al Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. *J Clin Endocrinol Metab*. 2000 May;85(5):1895-900. doi: 10.1210/jcem.85.5.6603. PMID: 10843171.
[l] McClung MR, Geusens P, Miller PD, et al Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. *N Engl J Med*. 2001 Feb 1;344(5):333-40. doi: 10.1056/nejm200102013440503. PMID: 11172164.
[m] Hosking D, Adami S, Felsenberg D, et al Comparison of change in bone resorption and bone mineral density with once-weekly alendronate and daily risedronate: a randomised, placebo-controlled study. *Curr Med Res Opin*. 2003;19(5):383-94. doi: 10.1185/030079903125002009. PMID: 13678475.
[n] Shiraki M, Fukunaga M, Kushida K, et al A double-blind dose-ranging study of risedronate in Japanese patients with osteoporosis (a study by the Risedronate Late Phase II Research Group). Osteoporos Int. 2003 May;14(3):225-34. doi: 10.1007/s00198-002-1369-9. PMID: 12730746.
[o] Valimaki MJ, Farrerons-Minguella J, Halse J, et al Effects of risedronate 5 mg/d on bone mineral density and bone turnover markers in late-postmenopausal women with osteopenia: a multinational, 24-month, randomized, double-blind, placebo-controlled, parallel-group, phase III trial. *Clin Ther*. 2007 Sep;29(9):1937-49. doi: 10.1016/j.clinthera.2007.09.017. PMID: 18035193.
[p] Reid IR, Brown JP, Burckhardt P, et al Intravenous zoledronic acid in postmenopausal women with low bone mineral density. *N Engl J Med*. 2002 Feb 28;346(9):653-61. doi: 10.1056/NEJMoa011807. PMID: 11870242.
[q] McClung M, Miller P, Recknor C, et al Zoledronic acid for the prevention of bone loss in postmenopausal women with low bone mass: a randomized controlled trial. Obstet Gynecol. 2009 Nov;114(5):999-1007. doi: 10.1097/AOG.0b013e3181bdce0a. PMID: 20168099.
[r] Boonen S, Reginster JY, Kaufman JM, et al Fracture risk and zoledronic acid therapy in men with osteoporosis. N Engl J Med. 2012 Nov;367(18):1714-23. doi: 10.1056/NEJMoa1204061. PMID: 23113482.

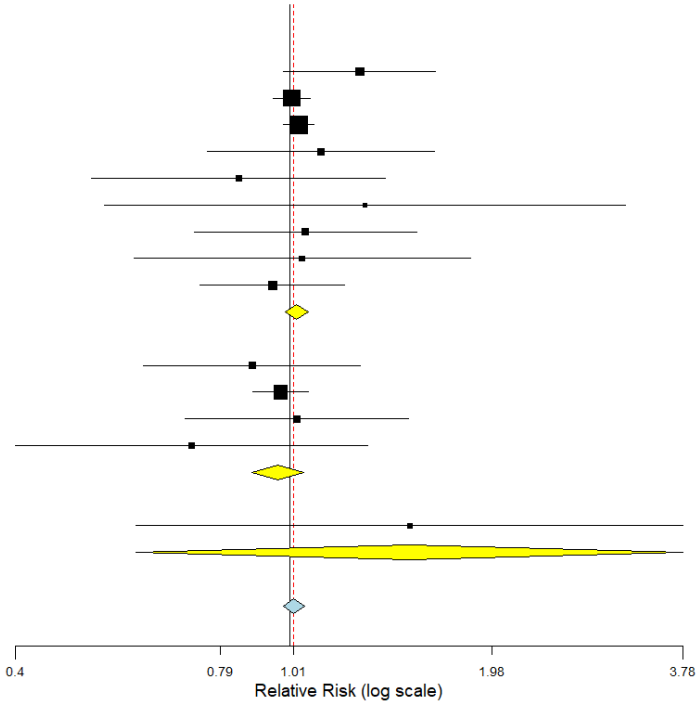
Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 27. Upper gastrointestinal events for bisphosphonates versus placebo (KQ5)

Studies	Relative risk (95% CI)	Events/Treatment	Events/Placebo
a. Tucci 1996	1.267 (0.980, 1.638)	49/94	79/192
b. Cummings 1998	1.007 (0.946, 1.071)	1052/2214	1047/2218
c. Bauer 2000	1.030 (0.978, 1.085)	1536/3226	1490/3223
d. Adachi 2001	1.111 (0.757, 1.630)	66/291	30/147
e. Greenspan 2002	0.841 (0.511, 1.383)	25/224	30/226
f. Ascott-Evans 2003	1.289 (0.534, 3.114)	15/95	6/49
g. Hosking (alendronate) 2003	1.054 (0.724, 1.535)	62/219	29/108
h. Eisman 2004	1.043 (0.591, 1.842)	22/225	21/224
i. Cryer 2005	0.943 (0.739, 1.204)	79/224	86/230
Subgroup Alendronate (I^2=0 % , P=0.812)	**1.024 (0.985, 1.064)**	**2906/6812**	**2818/6617**
j. Fogelman 2000	0.880 (0.610, 1.270)	40/174	47/180
k. McClung 2001	0.970 (0.882, 1.066)	657/3104	684/3134
l. Hosking (risedronate) 2003	1.023 (0.701, 1.493)	61/222	29/108
m. Valimaki 2007	0.717 (0.396, 1.301)	21/115	14/55
Subgroup Risedronate (I^2=0 % , P=0.732)	**0.961 (0.880, 1.049)**	**779/3615**	**774/3477**
n. Reginster 2005	1.500 (0.595, 3.779)	9/36	6/36
Subgroup Ibandronate (I^2=NA , P=NA)	**1.500 (0.595, 3.779)**	**9/36**	**6/36**
Overall (I^2=0 % , P=0.835)	**1.014 (0.979, 1.050)**	**3694/10463**	**3598/10130**



[a] Ascott-Evans BH, Guanabens N, Kivinen S, et al Alendronate prevents loss of bone density associated with discontinuation of hormone replacement therapy: a randomized controlled trial. Arch Intern Med. 2003 Apr 14;163(7):789-94. doi: 10.1001/archinte.163.7.789. PMID: 12695269.

[b] Tucci JR, Tonino RP, Emkey RD, et al Effect of three years of oral alendronate treatment in postmenopausal women with osteoporosis. *Am J Med*. 1996 Nov;101(5):488-501. PMID: 8948272.

[c] Cummings SR, Black DM, Thompson DE, et al Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. *JAMA*. 1998 Dec 23-30;280(24):2077-82. PMID: 9875874.

[d] Bauer DC, Black D, Ensrud K, et al Upper gastrointestinal tract safety profile of alendronate: the fracture intervention trial. Arch Intern Med. 2000 Feb 28;160(4):517-25. PMID: 10695692.

[e] Adachi JD, Saag KG, Delmas PD, et al Two-year effects of alendronate on bone mineral density and vertebral fracture in patients receiving glucocorticoids: a randomized, double-blind, placebo-controlled extension trial. Arthritis Rheum. 2001 Jan;44(1):202-11. doi: 10.1002/1529-0131(200101)44:1<202::aid-anr27>3.0.co;2-w. PMID: 11212161.

[f] Greenspan S, Field-Munves E, Tonino R, et al Tolerability of once-weekly alendronate in patients with osteoporosis: a randomized, double-blind, placebo-controlled study. *Mayo Clin Proc*. 2002 Oct;77(10):1044-52. doi: 10.4065/77.10.1044. PMID: 12374248.

[g] Hosking D, Adami S, Felsenberg D, et al Comparison of change in bone resorption and bone mineral density with once-weekly alendronate and daily risedronate: a randomised, placebo-controlled study. *Curr Med Res Opin*. 2003;19(5):383-94. doi: 10.1185/030079903125002009. PMID: 13678475.

© 2018 American Medical Association. All rights reserved.

[h] Eisman JA, Rizzoli R, Roman-Ivorra J, et al Upper gastrointestinal and overall tolerability of alendronate once weekly in patients with osteoporosis: results of a randomized, double-blind, placebo-controlled study. Curr Med Res Opin. 2004 May;20(5):699-705. doi: 10.1185/030079904125003548. PMID: 15140336.
[i] Cryer B, Binkley N, Simonelli C, et al A randomized, placebo-controlled, 6-month study of once-weekly alendronate oral solution for postmenopausal osteoporosis. Am J Geriatr Pharmacother. 2005 Sep;3(3):127-36. PMID: 16257815.
[j] Fogelman I, Ribot C, Smith R, et al Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. *J Clin Endocrinol Metab*. 2000 May;85(5):1895-900. doi: 10.1210/jcem.85.5.6603. PMID: 10843171.
[k] McClung MR, Geusens P, Miller PD, et al Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. *N Engl J Med*. 2001 Feb 1;344(5):333-40. doi: 10.1056/nejm200102013440503. PMID: 11172164.
[l] Hosking D, Adami S, Felsenberg D, et al Comparison of change in bone resorption and bone mineral density with once-weekly alendronate and daily risedronate: a randomised, placebo-controlled study. *Curr Med Res Opin*. 2003;19(5):383-94. doi: 10.1185/030079903125002009. PMID: 13678475.
[m] Valimaki MJ, Farrerons-Minguella J, Halse J, et al Effects of risedronate 5 mg/d on bone mineral density and bone turnover markers in late-postmenopausal women with osteopenia: a multinational, 24-month, randomized, double-blind, placebo-controlled, parallel-group, phase III trial. *Clin Ther*. 2007 Sep;29(9):1937-49. doi: 10.1016/j.clinthera.2007.09.017. PMID: 18035193.
[n] Reginster JY, Wilson KM, Dumont E, et al Monthly oral ibandronate is well tolerated and efficacious in postmenopausal women: results from the monthly oral pilot study. *J Clin Endocrinol Metab*. 2005 Sep;90(9):5018-24. doi: 10.1210/jc.2004-1750. PMID: 15972582.

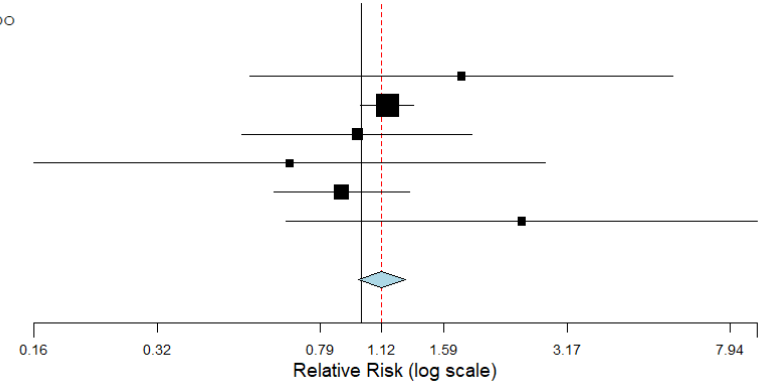
Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 28. Discontinuations due to adverse events for raloxifene versus placebo (KQ5)

Studies	Relative risk (95% CI)	Events/Treatment	Events/Placebo
a. Johnell 2002	1.750 (0.533, 5.751)	7/82	4/82
b. MORE 1999	1.156 (0.996, 1.342)	327/2557	285/2576
c. McClung 2006	0.976 (0.512, 1.862)	23/163	12/83
d. Meunier 1999	0.667 (0.159, 2.800)	3/45	4/40
e. Miller 2008	0.893 (0.610, 1.306)	43/311	48/310
f. Morii 2003	2.460 (0.656, 9.229)	7/92	3/97
Overall (I^2=0 % , P=0.533)	**1.121 (0.981, 1.281)**	**410/3250**	**356/3188**



[a] Johnell O, Scheele WH, Lu Y, et al. Additive effects of raloxifene and alendronate on bone density and biochemical markers of bone remodeling in postmenopausal women with osteoporosis. J Clin Endocrinol Metab. 2002 Mar;87(3):985-92. doi: 10.1210/jcem.87.3.8325. PMID: 11889149.

[b] Ettinger B, Black DM, Mitlak BH, et al. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. Multiple Outcomes of Raloxifene Evaluation (MORE) Investigators. JAMA. 1999 Aug 18;282(7):637-45. PMID: 10517716.; Delmas PD, Ensrud KE, Adachi JD, et al. Efficacy of raloxifene on vertebral fracture risk reduction in postmenopausal women with osteoporosis: four-year results from a randomized clinical trial. J Clin Endocrinol Metab. 2002 Aug;87(8):3609-17. doi: 10.1210/jcem.87.8.8750. PMID: 12161484; Barrett-Connor E, Cauley JA, Kulkarni PM, et al. Risk-benefit profile for raloxifene: 4-year data From the Multiple Outcomes of Raloxifene Evaluation (MORE) randomized trial. J Bone Miner Res. 2004 Aug;19(8):1270-5. doi: 10.1359/JBMR.040406 [doi]. PMID: 15231013; Barrett-Connor E, Grady D, Sashegyi A, et al. Raloxifene and cardiovascular events in osteoporotic postmenopausal women: four-year results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA. 2002 Feb 20;287(7):847-57. doi: joc11015 [pii]. PMID: 11851576; Keech CA, Sashegyi A, Barrett-Connor E. Year-by-year analysis of cardiovascular events in the Multiple Outcomes of Raloxifene Evaluation (MORE) trial. Curr Med Res Opin. 2005 Jan;21(1):135-40. PMID: 15881485; Cauley JA, Norton L, Lippman ME, et al. Continued breast cancer risk reduction in postmenopausal women treated with raloxifene: 4-year results from the MORE trial. Multiple outcomes of raloxifene evaluation. Breast Cancer Res Treat. 2001 Jan;65(2):125-34. PMID: 11261828.

[c] McClung MR, Siris E, Cummings S, et al. Prevention of bone loss in postmenopausal women treated with lasofoxifene compared with raloxifene. Menopause. 2006 May-Jun;13(3):377-86. doi: 10.1097/01.gme.0000188736.69617.4f. PMID: 16735934.

[d] Meunier PJ, Vignot E, Garnero P, et al. Treatment of postmenopausal women with osteoporosis or low bone density with raloxifene. Raloxifene Study Group. Osteoporos Int. 1999;10(4):330-6. PMID: 10692984.

[e] Miller PD, Chines AA, Christiansen C, et al. Effects of bazedoxifene on BMD and bone turnover in postmenopausal women: 2-yr results of a randomized, double-blind, placebo-, and active-controlled study. J Bone Miner Res. 2008 Apr;23(4):525-35. doi: 10.1359/jbmr.071206. PMID: 18072873.

[f] Morii H, Ohashi Y, Taketani Y, et al. Effect of raloxifene on bone mineral density and biochemical markers of bone turnover in Japanese postmenopausal women with osteoporosis: results from a randomized placebo-controlled trial. Osteoporos Int. 2003 Oct;14(10):793-800. doi: 10.1007/s00198-003-1424-1. PMID: 12955333.

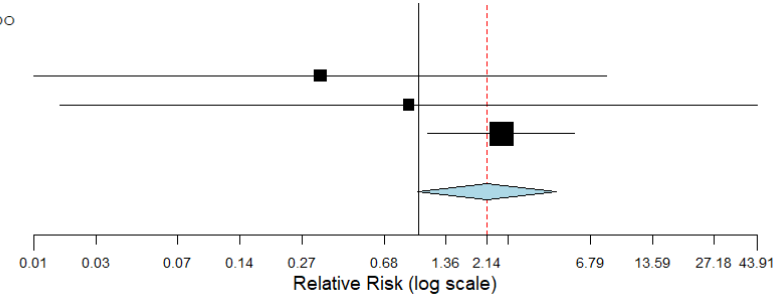
Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 29. Deep vein thrombosis for raloxifene versus placebo (KQ5)

Studies	Relative risk (95% CI)	Events/Treatment	Events/Placebo
a. Miller 2008	0.332 (0.014, 8.125)	0/311	1/310
b. Meunier 1999	0.891 (0.018, 43.912)	0/45	0/40
c. MORE 1999	2.519 (1.111, 5.707)	20/2557	8/2576
Overall (I^2=0 % , P=0.438)	**2.144 (0.986, 4.662)**	**20/2913**	**9/2926**



[a] Miller PD, Chines AA, Christiansen C, et al. Effects of bazedoxifene on BMD and bone turnover in postmenopausal women: 2-yr results of a randomized, double-blind, placebo-, and active-controlled study. J Bone Miner Res. 2008 Apr;23(4):525-35. doi: 10.1359/jbmr.071206. PMID: 18072873.

[b] Meunier PJ, Vignot E, Garnero P, et al. Treatment of postmenopausal women with osteoporosis or low bone density with raloxifene. Raloxifene Study Group. Osteoporos Int. 1999;10(4):330-6. PMID: 10692984.

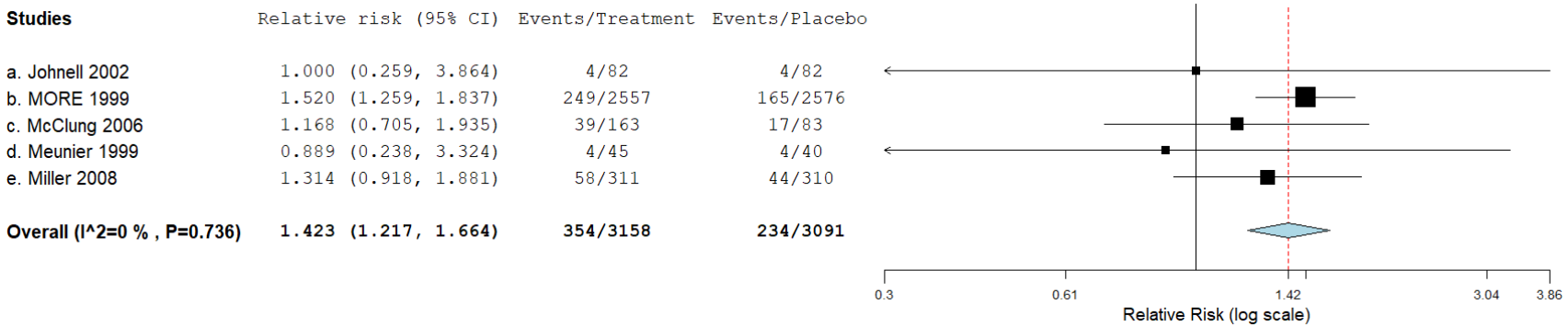
[c] Ettinger B, Black DM, Mitlak BH, et al. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. Multiple Outcomes of Raloxifene Evaluation (MORE) Investigators. JAMA. 1999 Aug 18;282(7):637-45. PMID: 10517716.; Delmas PD, Ensrud KE, Adachi JD, et al. Efficacy of raloxifene on vertebral fracture risk reduction in postmenopausal women with osteoporosis: four-year results from a randomized clinical trial. J Clin Endocrinol Metab. 2002 Aug;87(8):3609-17. doi: 10.1210/jcem.87.8.8750. PMID: 12161484; Barrett-Connor E, Cauley JA, Kulkarni PM, et al. Risk-benefit profile for raloxifene: 4-year data From the Multiple Outcomes of Raloxifene Evaluation (MORE) randomized trial. J Bone Miner Res. 2004 Aug;19(8):1270-5. doi: 10.1359/JBMR.040406 [doi]. PMID: 15231013; Barrett-Connor E, Grady D, Sashegyi A, et al. Raloxifene and cardiovascular events in osteoporotic postmenopausal women: four-year results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA. 2002 Feb 20;287(7):847-57. doi: joc11015 [pii]. PMID: 11851576; Keech CA, Sashegyi A, Barrett-Connor E. Year-by-year analysis of cardiovascular events in the Multiple Outcomes of Raloxifene Evaluation (MORE) trial. Curr Med Res Opin. 2005 Jan;21(1):135-40. PMID: 15881485; Cauley JA, Norton L, Lippman ME, et al. Continued breast cancer risk reduction in postmenopausal women treated with raloxifene: 4-year results from the MORE trial. Multiple outcomes of raloxifene evaluation. Breast Cancer Res Treat. 2001 Jan;65(2):125-34. PMID: 11261828.

Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 30. Hot flashes for raloxifene versus placebo (KQ5)



[a] Johnell O, Scheele WH, Lu Y, et al. Additive effects of raloxifene and alendronate on bone density and biochemical markers of bone remodeling in postmenopausal women with osteoporosis. J Clin Endocrinol Metab. 2002 Mar;87(3):985-92. doi: 10.1210/jcem.87.3.8325. PMID: 11889149.

[b] Ettinger B, Black DM, Mitlak BH, et al. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. Multiple Outcomes of Raloxifene Evaluation (MORE) Investigators. JAMA. 1999 Aug 18;282(7):637-45. PMID: 10517716.; Delmas PD, Ensrud KE, Adachi JD, et al. Efficacy of raloxifene on vertebral fracture risk reduction in postmenopausal women with osteoporosis: four-year results from a randomized clinical trial. J Clin Endocrinol Metab. 2002 Aug;87(8):3609-17. doi: 10.1210/jcem.87.8.8750. PMID: 12161484; Barrett-Connor E, Cauley JA, Kulkarni PM, et al. Risk-benefit profile for raloxifene: 4-year data From the Multiple Outcomes of Raloxifene Evaluation (MORE) randomized trial. J Bone Miner Res. 2004 Aug;19(8):1270-5. doi: 10.1359/JBMR.040406 [doi]. PMID: 15231013; Barrett-Connor E, Grady D, Sashegyi A, et al. Raloxifene and cardiovascular events in osteoporotic postmenopausal women: four-year results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA. 2002 Feb 20;287(7):847-57. doi: joc11015 [pii]. PMID: 11851576; Keech CA, Sashegyi A, Barrett-Connor E. Year-by-year analysis of cardiovascular events in the Multiple Outcomes of Raloxifene Evaluation (MORE) trial. Curr Med Res Opin. 2005 Jan;21(1):135-40. PMID: 15881485; Cauley JA, Norton L, Lippman ME, et al. Continued breast cancer risk reduction in postmenopausal women treated with raloxifene: 4-year results from the MORE trial. Multiple outcomes of raloxifene evaluation. Breast Cancer Res Treat. 2001 Jan;65(2):125-34. PMID: 11261828.

[c] McClung MR, Siris E, Cummings S, et al. Prevention of bone loss in postmenopausal women treated with lasofoxifene compared with raloxifene. Menopause. 2006 May-Jun;13(3):377-86. doi: 10.1097/01.gme.0000188736.69617.4f. PMID: 16735934.

[d] Meunier PJ, Vignot E, Garnero P, et al. Treatment of postmenopausal women with osteoporosis or low bone density with raloxifene. Raloxifene Study Group. Osteoporos Int. 1999;10(4):330-6. PMID: 10692984.

[e] Miller PD, Chines AA, Christiansen C, et al. Effects of bazedoxifene on BMD and bone turnover in postmenopausal women: 2-yr results of a randomized, double-blind, placebo-, and active-controlled study. J Bone Miner Res. 2008 Apr;23(4):525-35. doi: 10.1359/jbmr.071206. PMID: 18072873.

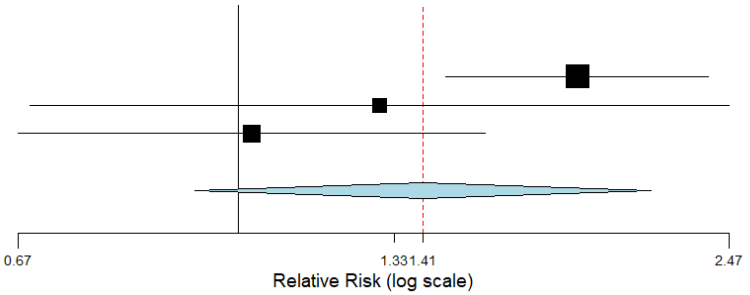
Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 31. Leg cramps for raloxifene versus placebo (KQ5)

Studies	Relative risk (95% CI)	Events/Treatment	Events/Placebo
a. MORE 1999	1.868 (1.466, 2.380)	178/2557	96/2576
b. McClung 2006	1.296 (0.680, 2.471)	28/163	11/83
c. Miller 2008	1.024 (0.666, 1.577)	37/311	36/310
Overall (I^2=67.1 % , P=0.048)	**1.406 (0.923, 2.142)**	**243/3031**	**143/2969**



[a] Ettinger B, Black DM, Mitlak BH, et al. Reduction of vertebral fracture risk in postmenopausal women with osteoporosis treated with raloxifene: results from a 3-year randomized clinical trial. Multiple Outcomes of Raloxifene Evaluation (MORE) Investigators. JAMA. 1999 Aug 18;282(7):637-45. PMID: 10517716.; Delmas PD, Ensrud KE, Adachi JD, et al. Efficacy of raloxifene on vertebral fracture risk reduction in postmenopausal women with osteoporosis: four-year results from a randomized clinical trial. J Clin Endocrinol Metab. 2002 Aug;87(8):3609-17. doi: 10.1210/jcem.87.8.8750. PMID: 12161484; Barrett-Connor E, Cauley JA, Kulkarni PM, et al. Risk-benefit profile for raloxifene: 4-year data From the Multiple Outcomes of Raloxifene Evaluation (MORE) randomized trial. J Bone Miner Res. 2004 Aug;19(8):1270-5. doi: 10.1359/JBMR.040406 [doi]. PMID: 15231013; Barrett-Connor E, Grady D, Sashegyi A, et al. Raloxifene and cardiovascular events in osteoporotic postmenopausal women: four-year results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA. 2002 Feb 20;287(7):847-57. doi: joc11015 [pii]. PMID: 11851576; Keech CA, Sashegyi A, Barrett-Connor E. Year-by-year analysis of cardiovascular events in the Multiple Outcomes of Raloxifene Evaluation (MORE) trial. Curr Med Res Opin. 2005 Jan;21(1):135-40. PMID: 15881485; Cauley JA, Norton L, Lippman ME, et al. Continued breast cancer risk reduction in postmenopausal women treated with raloxifene: 4-year results from the MORE trial. Multiple outcomes of raloxifene evaluation. Breast Cancer Res Treat. 2001 Jan;65(2):125-34. PMID: 11261828.

[b] McClung MR, Siris E, Cummings S, et al. Prevention of bone loss in postmenopausal women treated with lasofoxifene compared with raloxifene. Menopause.

[c] Miller PD, Chines AA, Christiansen C, et al. Effects of bazedoxifene on BMD and bone turnover in postmenopausal women: 2-yr results of a randomized, double-blind, placebo-, and active-controlled study. J Bone Miner Res. 2008 Apr;23(4):525-35. doi: 10.1359/jbmr.071206. PMID: 18072873.

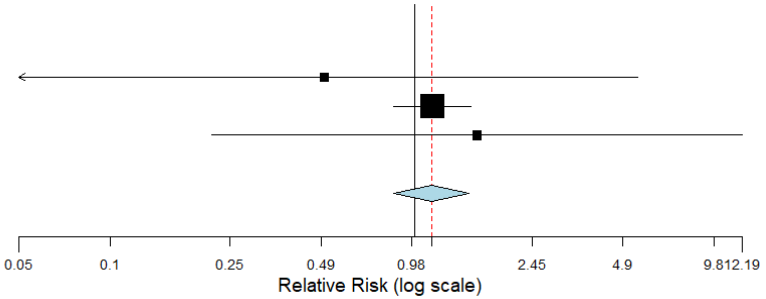
Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 32. Discontinuations due to adverse events for denosumab versus placebo (KQ5)

Studies	Relative risk (95% CI)	Events/Treatment	Events/Placebo
a. Bone 2008	0.503 (0.046, 5.494)	1/164	2/165
b. Cummings 2009	1.145 (0.853, 1.537)	93/3886	81/3876
c. Lewiecki 2007	1.611 (0.213, 12.192)	11/314	1/46
Overall (I^2=0 % , P=0.755)	**1.139 (0.853, 1.522)**	**105/4364**	**84/4087**



[a] Bone HG; Bolognese MA; Yuen CK; Kendler DL; Wang H; Liu Y; San Martin J. Effects of denosumab on bone mineral density and bone turnover in postmenopausal women. J Clin Endocrinol Metab. 2008 Jun;93(6):2149-57. doi: 10.1210/jc.2007-2814. PMID: 18381571.

[b] Cummings SR; San Martin J; McClung MR; Siris ES; Eastell R; Reid IR; Delmas P; Zoog HB; Austin M; Wang A; Kutilek S; Adami S; Zanchetta J; Libanati C; Siddhanti S; Christiansen C. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. N Engl J Med. 2009 Aug 20;361(8):756-65. doi: 10.1056/NEJMoa0809493. PMID: 19671655.

[c] Lewiecki EM; Miller PD; McClung MR; Cohen SB; Bolognese MA; Liu Y; Wang A; Siddhanti S; Fitzpatrick LA. Two-year treatment with denosumab (AMG 162) in a randomized phase 2 study of postmenopausal women with low BMD. J Bone Miner Res. 2007 Dec;22(12):1832-41. doi: 10.1359/jbmr.070809 [doi]. PMID: 17708711.

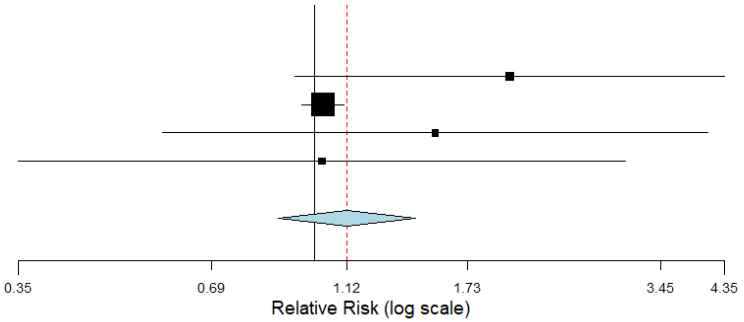
Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 33. Serious adverse events for denosumab versus placebo (KQ5)

Studies	Relative risk (95% CI)	Events/Treatment	Events/Placebo
a. Bone 2008	2.012 (0.931, 4.348)	18/164	9/165
b. Cummings 2009	1.030 (0.955, 1.112)	1004/3886	972/3876
c. Lewiecki 2007	1.538 (0.579, 4.089)	42/314	4/46
d. Nakamura 2011	1.025 (0.345, 3.045)	12/158	4/54
Overall (I^2=14.13 % , P=0.322)	**1.121 (0.876, 1.435)**	**1076/4522**	**989/4141**



[a] Bone HG; Bolognese MA; Yuen CK; Kendler DL; Wang H; Liu Y; San Martin J. Effects of denosumab on bone mineral density and bone turnover in postmenopausal women. J Clin Endocrinol Metab. 2008 Jun;93(6):2149-57. doi: 10.1210/jc.2007-2814. PMID: 18381571.

[b] Cummings SR; San Martin J; McClung MR; Siris ES; Eastell R; Reid IR; Delmas P; Zoog HB; Austin M; Wang A; Kutilek S; Adami S; Zanchetta J; Libanati C; Siddhanti S; Christiansen C. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. N Engl J Med. 2009 Aug 20;361(8):756-65. doi: 10.1056/NEJMoa0809493. PMID: 19671655.

[c] Lewiecki EM; Miller PD; McClung MR; Cohen SB; Bolognese MA; Liu Y; Wang A; Siddhanti S; Fitzpatrick LA. Two-year treatment with denosumab (AMG 162) in a randomized phase 2 study of postmenopausal women with low BMD. J Bone Miner Res. 2007 Dec;22(12):1832-41. doi: 10.1359/jbmr.070809 [doi]. PMID: 17708711.

[d] Nakamura T; Matsumoto T; Sugimoto T; Shiraki M. Dose-response study of denosumab on bone mineral density and bone turnover markers in Japanese postmenopausal women with osteoporosis. Osteoporos Int; 2012. p. 1131-40.

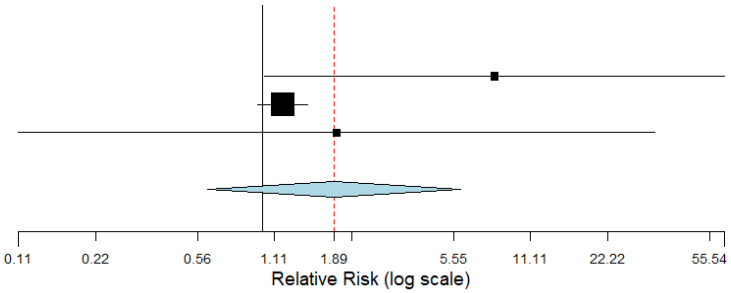
Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

eFigure 34. Serious infections for denosumab versus placebo (KQ5)

Studies	Relative risk (95% CI)	Events/Treatment	Events/Placebo
a. Bone 2008	8.049 (1.018, 63.632)	8/164	1/165
b. Cummings 2009	1.192 (0.951, 1.495)	159/3886	133/3876
c. Lewiecki 2007	1.940 (0.111, 33.872)	6/314	0/46
Overall (I^2=40.09 % , P=0.188)	**1.894 (0.607, 5.910)**	**173/4364**	**134/4087**



[a] Bone HG; Bolognese MA; Yuen CK; Kendler DL; Wang H; Liu Y; San Martin J. Effects of denosumab on bone mineral density and bone turnover in postmenopausal women. J Clin Endocrinol Metab. 2008 Jun;93(6):2149-57. doi: 10.1210/jc.2007-2814. PMID: 18381571.

[b] Cummings SR; San Martin J; McClung MR; Siris ES; Eastell R; Reid IR; Delmas P; Zoog HB; Austin M; Wang A; Kutilek S; Adami S; Zanchetta J; Libanati C; Siddhanti S; Christiansen C. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. N Engl J Med. 2009 Aug 20;361(8):756-65. doi: 10.1056/NEJMoa0809493. PMID: 19671655.

[c] Lewiecki EM; Miller PD; McClung MR; Cohen SB; Bolognese MA; Liu Y; Wang A; Siddhanti S; Fitzpatrick LA. Two-year treatment with denosumab (AMG 162) in a randomized phase 2 study of postmenopausal women with low BMD. J Bone Miner Res. 2007 Dec;22(12):1832-41. doi: 10.1359/jbmr.070809 [doi]. PMID: 17708711.

Abbreviations: AUC=confidence interval

The size of the boxes reflects the relative contribution of each study to the pooled estimate; the line surrounding the box represents the 95% confidence interval for each study's estimate of effect; the center of the diamond is the pooled estimate; and the width of the diamond represents 95% confidence intervals for the pooled estimate.

© 2018 American Medical Association. All rights reserved.

References

1. Barr RJ, Stewart A, Torgerson DJ, et al. Population screening for osteoporosis risk: a randomised control trial of medication use and fracture risk. Osteoporos Int. 2010 Apr;21(4):561-8. doi: 10.1007/s00198-009-1007-x [doi]. PMID: 19565176.
2. 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 Feb 24;391(10122):741-7. doi: 10.1016/s0140-6736(17)32640-5. PMID: 29254858.
3. Crandall CJ. Risk Assessment Tools for Osteoporosis Screening in Postmenopausal Women: A Systematic Review. Curr Osteoporos Rep. 2015 Oct;13(5):287-301. doi: 10.1007/s11914-015-0282-z. PMID: 26233285.
4. Marques A, Ferreira RJ, Santos E, et al. The accuracy of osteoporotic fracture risk prediction tools: a systematic review and meta-analysis. Ann Rheum Dis. 2015 Nov;74(11):1958-67. doi: 10.1136/annrheumdis-2015-207907. PMID: 26248637.
5. Nayak S, Edwards DL, Saleh AA, et al. Performance of risk assessment instruments for predicting osteoporotic fracture risk: a systematic review. Osteoporos Int. 2014 Jan;25(1):23-49. doi: 10.1007/s00198-013-2504-5. PMID: 24105431.
6. Rubin KH, Friis-Holmberg T, Hermann AP, et al. Risk assessment tools to identify women with increased risk of osteoporotic fracture: complexity or simplicity? A systematic review. J Bone Miner Res. 2013 Aug;28(8):1701-17. doi: 10.1002/jbmr.1956. PMID: 23592255.
7. Steurer J, Haller C, Hauselmann H, et al. Clinical value of prognostic instruments to identify patients with an increased risk for osteoporotic fractures: systematic review. PLoS One. 2011;6(5):e19994. doi: 10.1371/journal.pone.0019994. PMID: 21625596.
8. Adler RA, Tran MT, Petkov VI. Performance of the Osteoporosis Self-assessment Screening Tool for osteoporosis in American men. Mayo Clin Proc. 2003 Jun;78(6):723-7. doi: 10.4065/78.6.723. PMID: 12934782.
9. Bansal S, Pecina JL, Merry SP, et al. US Preventative Services Task Force FRAX threshold has a low sensitivity to detect osteoporosis in women ages 50-64 years. Osteoporos Int. 2015 Apr;26(4):1429-33. doi: 10.1007/s00198-015-3026-0. PMID: 25614141.
10. Ben Sedrine W, Devogelaer JP, Kaufman JM, et al. Evaluation of the simple calculated osteoporosis risk estimation (SCORE) in a sample of white women from Belgium. Bone. 2001 Oct;29(4):374-80. PMID: 11595621.
11. Brenneman SK, Lacroix AZ, Buist DS, et al. Evaluation of decision rules to identify postmenopausal women for intervention related to osteoporosis. Dis Manag. 2003 Fall;6(3):159-68. doi: 10.1089/109350703322425509. PMID: 14570384.
12. Cadarette SM, Jaglal SB, Murray TM, et al. Evaluation of decision rules for referring women for bone densitometry by dual-energy x-ray absorptiometry. JAMA. 2001 Jul 4;286(1):57-63. PMID: 11434827.
13. Cadarette SM, McIsaac WJ, Hawker GA, et al. The validity of decision rules for selecting women with primary osteoporosis for bone mineral density testing. Osteoporos Int. 2004 May;15(5):361-6. doi: 10.1007/s00198-003-1552-7. PMID: 14730421.

© 2018 American Medical Association. All rights reserved.

14. Cass AR, Shepherd AJ, Carlson CA. Osteoporosis risk assessment and ethnicity: validation and comparison of 2 clinical risk stratification instruments. J Gen Intern Med. 2006 Jun;21(6):630-5. doi: 10.1111/j.1525-1497.2006.00459.x. PMID: 16808748.
15. Cass AR, Shepherd AJ. Validation of the Male Osteoporosis Risk Estimation Score (MORES) in a primary care setting. J Am Board Fam Med. 2013 Jul-Aug;26(4):436-44. doi: 10.3122/jabfm.2013.04.120182. PMID: 23833159.
16. 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 Jul;14(4):365-9. doi: 10.1370/afm.1945. PMID: 27401426.
17. Chan SP, Teo CC, Ng SA, et al. Validation of various osteoporosis risk indices in elderly Chinese females in Singapore. Osteoporos Int. 2006;17(8):1182-8. doi: 10.1007/s00198-005-0051-4 [doi]. PMID: 16699739.
18. Cook RB, Collins D, Tucker J, et al. Comparison of questionnaire and quantitative ultrasound techniques as screening tools for DXA. Osteoporos Int. 2005 Dec;16(12):1565-75. doi: 10.1007/s00198-005-1864-x [doi]. PMID: 15883661.
19. Crandall CJ, Larson J, Gourlay ML, et al. Osteoporosis screening in postmenopausal women 50 to 64 years old: comparison of US Preventive Services Task Force strategy and two traditional strategies in the Women's Health Initiative. J Bone Miner Res. 2014 Jul;29(7):1661-6. doi: 10.1002/jbmr.2174 [doi]. PMID: 24431262.
20. D'Amelio P, Tamone C, Pluviano F, et al. Effects of lifestyle and risk factors on bone mineral density in a cohort of Italian women: suggestion for a new decision rule. Calcif Tissue Int. 2005 Aug;77(2):72-8. doi: 10.1007/s00223-004-0253-3. PMID: 16059776.
21. D'Amelio P, Spertino E, Martino F, et al. Prevalence of postmenopausal osteoporosis in Italy and validation of decision rules for referring women for bone densitometry. Calcif Tissue Int. 2013;92(5):437-43.
22. Geusens P, Hochberg MC, van der Voort DJ, et al. Performance of risk indices for identifying low bone density in postmenopausal women. Mayo Clin Proc. 2002 Jul;77(7):629-37. PMID: 12108600.
23. Gnudi S, Sitta E. Clinical risk factor evaluation to defer postmenopausal women from bone mineral density measurement: an Italian study. J Clin Densitom. 2005 Summer;8(2):199-205. PMID: 15908708.
24. Gourlay ML, Miller WC, Richy F, et al. Performance of osteoporosis risk assessment tools in postmenopausal women aged 45-64 years. Osteoporos Int. 2005 Aug;16(8):921-7. doi: 10.1007/s00198-004-1775-2 [doi]. PMID: 16028108.
25. Gourlay ML, Powers JM, Lui LY, et al. Clinical performance of osteoporosis risk assessment tools in women aged 67 years and older. Osteoporos Int. 2008 Aug;19(8):1175-83. doi: 10.1007/s00198-007-0555-1. PMID: 18219434.
26. Harrison EJ, Adams JE. Application of a triage approach to peripheral bone densitometry reduces the requirement for central DXA but is not cost effective. Calcif Tissue Int. 2006 Oct;79(4):199-206. doi: 10.1007/s00223-005-0302-6. PMID: 16969598.
27. Jimenez-Nunez FG, Manrique-Arija S, Urena-Garnica I, et al. Reducing the need for central dual-energy X-ray absorptiometry in postmenopausal women: efficacy of a clinical algorithm including peripheral densitometry. Calcif Tissue Int. 2013 Jul;93(1):62-8. doi: 10.1007/s00223-013-9728-4 [doi]. PMID: 23608922.
28. Kung AW, Ho AY, Sedrine WB, et al. Comparison of a simple clinical risk index and quantitative bone ultrasound for identifying women at increased risk of osteoporosis.

© 2018 American Medical Association. All rights reserved.

Osteoporos Int. 2003 Sep;14(9):716-21. doi: 10.1007/s00198-003-1428-x [doi]. PMID: 12897978.
29. Kung AW, Ho AY, Ross PD, et al. Development of a clinical assessment tool in identifying Asian men with low bone mineral density and comparison of its usefulness to quantitative bone ultrasound. Osteoporos Int. 2005 Jul;16(7):849-55. doi: 10.1007/s00198-004-1778-z [doi]. PMID: 15611839.
30. Leslie WD, Lix LM, Johansson H, et al. Selection of women aged 50-64 yr for bone density measurement. J Clin Densitom. 2013 Oct-Dec;16(4):570-8. doi: 10.1016/j.jocd.2013.01.004. PMID: 23452870.
31. Lynn HS, Woo J, Leung PC, et al. An evaluation of osteoporosis screening tools for the osteoporotic fractures in men (MrOS) study. Osteoporos Int. 2008 Jul;19(7):1087-92. doi: 10.1007/s00198-007-0553-3. PMID: 18239959.
32. Machado P, Coutinho M, da Silva JA. Selecting men for bone densitometry: performance of osteoporosis risk assessment tools in Portuguese men. Osteoporos Int. 2010 Jun;21(6):977-83. doi: 10.1007/s00198-009-1036-5 [doi]. PMID: 19727909.
33. Martinez-Aguila D, Gomez-Vaquero C, Rozadilla A, et al. Decision rules for selecting women for bone mineral density testing: application in postmenopausal women referred to a bone densitometry unit. J Rheumatol. 2007 Jun;34(6):1307-12. PMID: 17552058.
34. Mauck KF, Cuddihy MT, Atkinson EJ, et al. Use of clinical prediction rules in detecting osteoporosis in a population-based sample of postmenopausal women. Arch Intern Med. 2005 Mar 14;165(5):530-6. doi: 10.1001/archinte.165.5.530. PMID: 15767529.
35. McLeod KM, Johnson S, Rasali D, et al. Discriminatory performance of the calcaneal quantitative ultrasound and osteoporosis self-assessment tool to select older women for dual-energy x-ray absorptiometry. J Clin Densitom. 2015 Apr-Jun;18(2):157-64. doi: 10.1016/j.jocd.2015.02.006. PMID: 25937306.
36. Morin S, Tsang JF, Leslie WD. Weight and body mass index predict bone mineral density and fractures in women aged 40 to 59 years. Osteoporos Int. 2009 Mar;20(3):363-70. doi: 10.1007/s00198-008-0688-x [doi]. PMID: 18633665.
37. Nguyen TV, Center JR, Pocock NA, et al. Limited utility of clinical indices for the prediction of symptomatic fracture risk in postmenopausal women. Osteoporos Int. 2004 Jan;15(1):49-55. doi: 10.1007/s00198-003-1511-3. PMID: 14593453.
38. Oh SM, Nam BH, Rhee Y, et al. Development and validation of osteoporosis risk-assessment model for Korean postmenopausal women. J Bone Miner Metab. 2013 Jul;31(4):423-32. doi: 10.1007/s00774-013-0426-0 [doi]. PMID: 23420298.
39. Oh SM, Song BM, Nam BH, et al. Development and validation of osteoporosis risk-assessment model for Korean men. Yonsei Med J. 2016 Jan;57(1):187-96. doi: 10.3349/ymj.2016.57.1.187. PMID: 26632400.
40. Pang WY, Inderjeeth CA. FRAX without bone mineral density versus osteoporosis self-assessment screening tool as predictors of osteoporosis in primary screening of individuals aged 70 and older. J Am Geriatr Soc. 2014 Mar;62(3):442-6. doi: 10.1111/jgs.12696 [doi]. PMID: 24617899.
41. 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.
42. Richy F, Gourlay M, Ross PD, et al. Validation and comparative evaluation of the osteoporosis self-assessment tool (OST) in a Caucasian population from Belgium. QJM. 2004 Jan;97(1):39-46. PMID: 14702510.

© 2018 American Medical Association. All rights reserved.

43. Shepherd AJ, Cass AR, Carlson CA, et al. Development and internal validation of the male osteoporosis risk estimation score. Ann Fam Med. 2007 Nov-Dec;5(6):540-6. doi: 10.1370/afm.753. PMID: 18025492.
44. 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 Mar-Apr;23(2):186-94. doi: 10.3122/jabfm.2010.02.090027. PMID: 20207929.
45. 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. doi: 10.1007/s00198-005-0034-5. PMID: 16523248.
46. 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 Jul-Sep;10(3):289-97. doi: 10.1016/j.jocd.2007.03.001. PMID: 17459748.
47. Park HM, Sedrine WB, Reginster JY, et al. Korean experience with the OSTA risk index for osteoporosis: a validation study. J Clin Densitom. 2003 Fall;6(3):247-50. PMID: 14514994.
48. Boonen S, Nijs J, Borghs H, et al. Identifying postmenopausal women with osteoporosis by calcaneal ultrasound, metacarpal digital X-ray radiogrammetry and phalangeal radiographic absorptiometry: a comparative study. Osteoporos Int. 2005 Jan;16(1):93-100. doi: 10.1007/s00198-004-1660-z [doi]. PMID: 15197540.
49. Minnock E, Cook R, Collins D, et al. Using risk factors and quantitative ultrasound to identify postmenopausal caucasian women at risk of osteoporosis. J Clin Densitom. 2008 Oct-Dec;11(4):485-93. doi: 10.1016/j.jocd.2008.04.002. PMID: 18539491.
50. Richy F, Deceulaer F, Ethgen O, et al. Development and validation of the ORACLE score to predict risk of osteoporosis. Mayo Clin Proc. 2004 Nov;79(11):1402-8. doi: 10.4065/79.11.1402. PMID: 15544019.
51. Ahmed LA, Nguyen ND, Bjornerem A, et al. External validation of the Garvan nomograms for predicting absolute fracture risk: the Tromso study. PLoS One. 2014;9(9):e107695. doi: 10.1371/journal.pone.0107695. PMID: 25255221.
52. Azagra R, Roca G, Encabo G, et al. Prediction of absolute risk of fragility fracture at 10 years in a Spanish population: validation of the WHO FRAX tool in Spain. BMC Musculoskelet Disord. 2011;12:30. doi: 10.1186/1471-2474-12-30. PMID: 21272372.
53. Bauer DC, Ewing SK, Cauley JA, et al. Quantitative ultrasound predicts hip and non-spine fracture in men: the MrOS study. Osteoporos Int. 2007 Jun;18(6):771-7. doi: 10.1007/s00198-006-0317-5 [doi]. PMID: 17273893.
54. Berry SD, Samelson EJ, Pencina MJ, et al. Repeat bone mineral density screening and prediction of hip and major osteoporotic fracture. JAMA. 2013 Sep 25;310(12):1256-62. doi: 10.1001/jama.2013.277817. PMID: 24065012.
55. Chan MY, Nguyen ND, Center JR, et al. Absolute fracture-risk prediction by a combination of calcaneal quantitative ultrasound and bone mineral density. Calcif Tissue Int. 2012 Feb;90(2):128-36. doi: 10.1007/s00223-011-9556-3 [doi]. PMID: 22179560.
56. Chan MY, Nguyen ND, Center JR, et al. Quantitative ultrasound and fracture risk prediction in non-osteoporotic men and women as defined by WHO criteria. Osteoporos Int. 2013 Mar;24(3):1015-22. doi: 10.1007/s00198-012-2001-2 [doi]. PMID: 22878531.
57. Crandall CJ, Larson JC, Watts NB, et al. Comparison of fracture risk prediction by the US preventive services task force strategy and two alternative strategies in women 50-64

© 2018 American Medical Association. All rights reserved.

years old in the women's health initiative. J Clin Endocrinol Metab. 2014;99(12):4514-22.
58. Hans D, Goertzen AL, Krieg MA, et al. Bone microarchitecture assessed by TBS predicts osteoporotic fractures independent of bone density: the Manitoba study. J Bone Miner Res. 2011 Nov;26(11):2762-9. doi: 10.1002/jbmr.499 [doi]. PMID: 21887701.
59. Hillier TA, Stone KL, Bauer DC, et al. Evaluating the value of repeat bone mineral density measurement and prediction of fractures in older women: the study of osteoporotic fractures. Arch Intern Med. 2007 Jan 22;167(2):155-60. doi: 10.1001/archinte.167.2.155. PMID: 17242316.
60. 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. PMID: 22619194.
61. Iki M, Tamaki J, Kadowaki E, et al. Trabecular bone score (TBS) predicts vertebral fractures in Japanese women over 10 years independently of bone density and prevalent vertebral deformity: the Japanese Population-Based Osteoporosis (JPOS) cohort study. J Bone Miner Res. 2014 Feb;29(2):399-407. doi: 10.1002/jbmr.2048 [doi]. PMID: 23873699.
62. Iki M, Fujita Y, Tamaki J, et al. Trabecular bone score may improve FRAX(R) prediction accuracy for major osteoporotic fractures in elderly Japanese men: the Fujiwara-kyo Osteoporosis Risk in Men (FORMEN) Cohort Study. Osteoporos Int. 2015 Jun;26(6):1841-8. doi: 10.1007/s00198-015-3092-3. PMID: 25752623.
63. Kalvesten J, Lui LY, Brismar T, et al. Digital X-ray radiogrammetry in the study of osteoporotic fractures: Comparison to dual energy X-ray absorptiometry and FRAX. Bone. 2016:30-5. doi: 10.1016/j.bone.2016.02.011. PMID: CN-01137905.
64. Kanis JA, Oden A, Johnell O, et al. The use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. Osteoporos Int. 2007 Aug;18(8):1033-46. doi: 10.1007/s00198-007-0343-y. PMID: 17323110.
65. Kwok T, Khoo CC, Leung J, et al. Predictive values of calcaneal quantitative ultrasound and dual energy X ray absorptiometry for non-vertebral fracture in older men: results from the MrOS study (Hong Kong). Osteoporos Int. 2012 Mar;23(3):1001-6. doi: 10.1007/s00198-011-1634-x [doi]. PMID: 21528361.
66. Leslie WD, Lix LM. Simplified 10-year absolute fracture risk assessment: a comparison of men and women. J Clin Densitom. 2010 Apr-Jun;13(2):141-6. doi: 10.1016/j.jocd.2010.02.002. PMID: 20435264.
67. Leslie WD, Lix LM, Johansson H, et al. A comparative study of using non-hip bone density inputs with FRAX(R). Osteoporos Int. 2012 Mar;23(3):853-60. doi: 10.1007/s00198-011-1814-8 [doi]. PMID: 22008881.
68. Leslie WD, Morin S, Lix LM, et al. Fracture risk assessment without bone density measurement in routine clinical practice. Osteoporos Int. 2012 Jan;23(1):75-85. doi: 10.1007/s00198-011-1747-2. PMID: 21850546.
69. Leslie WD, Aubry-Rozier B, Lamy O, et al. TBS (trabecular bone score) and diabetes-related fracture risk. J Clin Endocrinol Metab. 2013 Feb;98(2):602-9. doi: 10.1210/jc.2012-3118. PMID: 23341489.

© 2018 American Medical Association. All rights reserved.

70. Lo JC, Pressman AR, Chandra M, et al. Fracture risk tool validation in an integrated healthcare delivery system. Am J Manag Care. 2011 Mar;17(3):188-94. doi: 48148 [pii]. PMID: 21504255.
71. Lundin H, Torabi F, Saaf M, et al. Laser-supported dual energy x-ray absorptiometry (DXL) compared to conventional absorptiometry (DXA) and to FRAX as tools for fracture risk assessments. PLoS One. 2015;10(9):e0137535. doi: 10.1371/journal.pone.0137535. PMID: 26413715.
72. Melton LJ, 3rd, Atkinson EJ, Khosla S, et al. Evaluation of a prediction model for long-term fracture risk. J Bone Miner Res. 2005 Apr;20(4):551-6. doi: 10.1359/JBMR.041206 [doi]. PMID: 15765172.
73. Miller PD, Siris ES, Barrett-Connor E, et al. Prediction of fracture risk in postmenopausal white women with peripheral bone densitometry: evidence from the National Osteoporosis Risk Assessment. J Bone Miner Res. 2002 Dec;17(12):2222-30. doi: 10.1359/jbmr.2002.17.12.2222 [doi]. PMID: 12469916.
74. Nguyen TV, Center JR, Eisman JA. Bone mineral density-independent association of quantitative ultrasound measurements and fracture risk in women. Osteoporos Int. 2004 Dec;15(12):942-7. doi: 10.1007/s00198-004-1717-z [doi]. PMID: 15309384.
75. Rubin KH, Abrahamsen B, Friis-Holmberg T, et al. Comparison of different screening tools (FRAX(R), OST, ORAI, OSIRIS, SCORE and age alone) to identify women with increased risk of fracture. A population-based prospective study. Bone. 2013 Sep;56(1):16-22. doi: 10.1016/j.bone.2013.05.002. PMID: 23669650.
76. Stewart A, Kumar V, Reid DM. Long-term fracture prediction by DXA and QUS: a 10-year prospective study. J Bone Miner Res. 2006 Mar;21(3):413-8. doi: 10.1359/jbmr.051205. PMID: 16491289.
77. van Geel TA, Eisman JA, Geusens PP, et al. The utility of absolute risk prediction using FRAX(R) and Garvan Fracture Risk Calculator in daily practice. Maturitas. 2014 Feb;77(2):174-9. doi: 10.1016/j.maturitas.2013.10.021. PMID: 24287178.
78. Crandall CJ, Newberry SJ, Diamant A, et al. Treatment To Prevent Fractures in Men and Women With Low Bone Density or Osteoporosis: Update of a 2007 Report Comparative Effectiveness Review No. 53. (Prepared by Southern California Evidence-based Practice Center under Contract No. HHSA-290-2007-10062-I.). AHRQ Publication No. 12-EHC023-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2012.
79. Abrahamsen B, Eiken P, Eastell R. Cumulative alendronate dose and the long-term absolute risk of subtrochanteric and diaphyseal femur fractures: a register-based national cohort analysis. J Clin Endocrinol Metab. 2010 Dec;95(12):5258-65. doi: 10.1210/jc.2010-1571. PMID: 20843943.
80. Adachi JD, Faraawi RY, O'Mahony MF, et al. Upper gastrointestinal tolerability of alendronate sodium monohydrate 10 mg once daily in postmenopausal women: a 12-week, randomized, double-blind, placebo-controlled, exploratory study. Clin Ther. 2009 Aug;31(8):1747-53. doi: 10.1016/j.clinthera.2009.08.016. PMID: 19808133.
81. Barrett-Connor E, Grady D, Sashegyi A, et al. Raloxifene and cardiovascular events in osteoporotic postmenopausal women: four-year results from the MORE (Multiple Outcomes of Raloxifene Evaluation) randomized trial. JAMA. 2002 Feb 20;287(7):847-57. doi: joc11015 [pii]. PMID: 11851576.
82. Barrett-Connor E, Cauley JA, Kulkarni PM, et al. Risk-benefit profile for raloxifene: 4-year data From the Multiple Outcomes of Raloxifene Evaluation (MORE) randomized

© 2018 American Medical Association. All rights reserved.

trial. J Bone Miner Res. 2004 Aug;19(8):1270-5. doi: 10.1359/JBMR.040406 [doi]. PMID: 15231013.
83. Bone HG, Greenspan SL, McKeever C, et al. Alendronate and estrogen effects in postmenopausal women with low bone mineral density. Alendronate/Estrogen Study Group. J Clin Endocrinol Metab. 2000 Feb;85(2):720-6. doi: 10.1210/jcem.85.2.6393. PMID: 10690882.
84. Bone HG, Bolognese MA, Yuen CK, et al. Effects of denosumab on bone mineral density and bone turnover in postmenopausal women. J Clin Endocrinol Metab. 2008 Jun;93(6):2149-57. doi: 10.1210/jc.2007-2814. PMID: 18381571.
85. Boonen S, Reginster JY, Kaufman JM, et al. Fracture risk and zoledronic acid therapy in men with osteoporosis. N Engl J Med. 2012 Nov;367(18):1714-23. doi: 10.1056/NEJMoa1204061. PMID: 23113482.
86. Cartsos VM, Zhu S, Zavras AI. Bisphosphonate use and the risk of adverse jaw outcomes: a medical claims study of 714,217 people. J Am Dent Assoc. 2008 Jan;139(1):23-30. PMID: 18167381.
87. Chapurlat RD, Laroche M, Thomas T, et al. Effect of oral monthly ibandronate on bone microarchitecture in women with osteopenia-a randomized placebo-controlled trial. Osteoporos Int. 2013 Jan;24(1):311-20. doi: 10.1007/s00198-012-1947-4 [doi]. PMID: 22402673.
88. Cryer B, Binkley N, Simonelli C, et al. A randomized, placebo-controlled, 6-month study of once-weekly alendronate oral solution for postmenopausal osteoporosis. Am J Geriatr Pharmacother. 2005 Sep;3(3):127-36. PMID: 16257815.
89. Cummings SR, Black DM, Thompson DE, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. JAMA. 1998 Dec 23-30;280(24):2077-82. PMID: 9875874.
90. Quandt SA, Thompson DE, Schneider DL, et al. Effect of alendronate on vertebral fracture risk in women with bone mineral density T scores of-1.6 to -2.5 at the femoral neck: the Fracture Intervention Trial. Mayo Clin Proc. 2005 Mar;80(3):343-9. PMID: 15757015.
91. Bauer DC, Black D, Ensrud K, et al. Upper gastrointestinal tract safety profile of alendronate: the fracture intervention trial. Arch Intern Med. 2000 Feb 28;160(4):517-25. PMID: 10695692.
92. Cummings SR, San Martin J, McClung MR, et al. Denosumab for prevention of fractures in postmenopausal women with osteoporosis. N Engl J Med. 2009 Aug 20;361(8):756-65. doi: 10.1056/NEJMoa0809493. PMID: 19671655.
93. Watts NB, Roux C, Modlin JF, et al. Infections in postmenopausal women with osteoporosis treated with denosumab or placebo: coincidence or causal association? Osteoporos Int. 2012 Jan;23(1):327-37. doi: 10.1007/s00198-011-1755-2 [doi]. PMID: 21892677.
94. McClung MR, Boonen S, Torring O, et al. Effect of denosumab treatment on the risk of fractures in subgroups of women with postmenopausal osteoporosis. J Bone Miner Res. 2012 Jan;27(1):211-8. doi: 10.1002/jbmr.536 [doi]. PMID: 21976367.
95. Boonen S, Adachi JD, Man Z, et al. Treatment with denosumab reduces the incidence of new vertebral and hip fractures in postmenopausal women at high risk. J Clin Endocrinol Metab. 2011 Jun;96(6):1727-36. doi: 10.1210/jc.2010-2784. PMID: 21411557.

© 2018 American Medical Association. All rights reserved.

96. Eisman JA, Rizzoli R, Roman-Ivorra J, et al. Upper gastrointestinal and overall tolerability of alendronate once weekly in patients with osteoporosis: results of a randomized, double-blind, placebo-controlled study. Curr Med Res Opin. 2004 May;20(5):699-705. doi: 10.1185/030079904125003548. PMID: 15140336.
97. Fogelman I, Ribot C, Smith R, et al. Risedronate reverses bone loss in postmenopausal women with low bone mass: results from a multinational, double-blind, placebo-controlled trial. BMD-MN Study Group. J Clin Endocrinol Metab. 2000 May;85(5):1895-900. doi: 10.1210/jcem.85.5.6603. PMID: 10843171.
98. Greenspan S, Field-Munves E, Tonino R, et al. Tolerability of once-weekly alendronate in patients with osteoporosis: a randomized, double-blind, placebo-controlled study. Mayo Clin Proc. 2002 Oct;77(10):1044-52. doi: 10.4065/77.10.1044. PMID: 12374248.
99. Greenspan SL, Resnick NM, Parker RA. Combination therapy with hormone replacement and alendronate for prevention of bone loss in elderly women: a randomized controlled trial. JAMA. 2003 May 21;289(19):2525-33. doi: 10.1001/jama.289.19.2525. PMID: 12759324.
100. Grey A, Bolland M, Wattie D, et al. Prolonged antiresorptive activity of zoledronate: a randomized, controlled trial. J Bone Miner Res. 2010 Oct;25(10):2251-5. doi: 10.1002/jbmr.103. PMID: 20499349.
101. Hosking D, Adami S, Felsenberg D, et al. Comparison of change in bone resorption and bone mineral density with once-weekly alendronate and daily risedronate: a randomised, placebo-controlled study. Curr Med Res Opin. 2003;19(5):383-94. doi: 10.1185/030079903125002009. PMID: 13678475.
102. Johnell O, Scheele WH, Lu Y, et al. Additive effects of raloxifene and alendronate on bone density and biochemical markers of bone remodeling in postmenopausal women with osteoporosis. J Clin Endocrinol Metab. 2002 Mar;87(3):985-92. doi: 10.1210/jcem.87.3.8325. PMID: 11889149.
103. Keech CA, Sashegyi A, Barrett-Connor E. Year-by-year analysis of cardiovascular events in the Multiple Outcomes of Raloxifene Evaluation (MORE) trial. Curr Med Res Opin. 2005 Jan;21(1):135-40. PMID: 15881485.
104. Kung AW, Yeung SS, Chu LW. The efficacy and tolerability of alendronate in postmenopausal osteoporotic Chinese women: a randomized placebo-controlled study. Calcif Tissue Int. 2000 Oct;67(4):286-90. PMID: 11000341.
105. Lasco A, Catalano A, Morabito N, et al. Adrenal effects of teriparatide in the treatment of severe postmenopausal osteoporosis. Osteoporos Int. 2011 Jan;22(1):299-303. doi: 10.1007/s00198-010-1222-5 [doi]. PMID: 20309523.
106. Lewiecki EM, Miller PD, McClung MR, et al. Two-year treatment with denosumab (AMG 162) in a randomized phase 2 study of postmenopausal women with low BMD. J Bone Miner Res. 2007 Dec;22(12):1832-41. doi: 10.1359/jbmr.070809 [doi]. PMID: 17708711.
107. McCloskey EV, Johansson H, Oden A, et al. Denosumab reduces the risk of osteoporotic fractures in postmenopausal women, particularly in those with moderate to high fracture risk as assessed with FRAX. J Bone Miner Res. 2012 Jul;27(7):1480-6. doi: 10.1002/jbmr.1606 [doi]. PMID: 22431426.
108. McClung MR, Wasnich RD, Recker R, et al. Oral daily ibandronate prevents bone loss in early postmenopausal women without osteoporosis. J Bone Miner Res. 2004 Jan;19(1):11-8. doi: 10.1359/jbmr.0301202. PMID: 14753731.

© 2018 American Medical Association. All rights reserved.

109. McClung MR, Siris E, Cummings S, et al. Prevention of bone loss in postmenopausal women treated with lasofoxifene compared with raloxifene. Menopause. 2006 May-Jun;13(3):377-86. doi: 10.1097/01.gme.0000188736.69617.4f. PMID: 16735934.
110. McClung MR, Lewiecki EM, Cohen SB, et al. Denosumab in postmenopausal women with low bone mineral density. N Engl J Med. 2006 Feb 23;354(8):821-31. doi: 10.1056/NEJMoa044459. PMID: 16495394.
111. McClung M, Miller P, Recknor C, et al. Zoledronic acid for the prevention of bone loss in postmenopausal women with low bone mass: a randomized controlled trial. Obstet Gynecol. 2009 Nov;114(5):999-1007. doi: 10.1097/AOG.0b013e3181bdce0a. PMID: 20168099.
112. Meunier PJ, Vignot E, Garnero P, et al. Treatment of postmenopausal women with osteoporosis or low bone density with raloxifene. Raloxifene Study Group. Osteoporos Int. 1999;10(4):330-6. PMID: 10692984.
113. Miller PD, Chines AA, Christiansen C, et al. Effects of bazedoxifene on BMD and bone turnover in postmenopausal women: 2-yr results of a randomized, double-blind, placebo-, and active-controlled study. J Bone Miner Res. 2008 Apr;23(4):525-35. doi: 10.1359/jbmr.071206. PMID: 18072873.
114. Morii H, Ohashi Y, Taketani Y, et al. Effect of raloxifene on bone mineral density and biochemical markers of bone turnover in Japanese postmenopausal women with osteoporosis: results from a randomized placebo-controlled trial. Osteoporos Int. 2003 Oct;14(10):793-800. doi: 10.1007/s00198-003-1424-1. PMID: 12955333.
115. Murphy MG, Weiss S, McClung M, et al. Effect of alendronate and MK-677 (a growth hormone secretagogue), individually and in combination, on markers of bone turnover and bone mineral density in postmenopausal osteoporotic women. J Clin Endocrinol Metab. 2001 Mar;86(3):1116-25. doi: 10.1210/jcem.86.3.7294. PMID: 11238495.
116. Nakamura T, Matsumoto T, Sugimoto T, et al. Dose-response study of denosumab on bone mineral density and bone turnover markers in Japanese postmenopausal women with osteoporosis. In Osteoporos Int
117. Orwoll ES, Scheele WH, Paul S, et al. The effect of teriparatide [human parathyroid hormone (1-34)] therapy on bone density in men with osteoporosis. J Bone Miner Res. 2003 Jan;18(1):9-17. doi: 10.1359/jbmr.2003.18.1.9. PMID: 12510800.
118. Pazianas M, Blumentals WA, Miller PD. Lack of association between oral bisphosphonates and osteonecrosis using jaw surgery as a surrogate marker. Osteoporos Int. 2008 Jun;19(6):773-9. doi: 10.1007/s00198-007-0547-1. PMID: 17999023.
119. Ravn P, Clemmesen B, Riis BJ, et al. The effect on bone mass and bone markers of different doses of ibandronate: a new bisphosphonate for prevention and treatment of postmenopausal osteoporosis: a 1-year, randomized, double-blind, placebo-controlled dose-finding study. Bone. 1996 Nov;19(5):527-33. PMID: 8922653.
120. Reginster JY, Wilson KM, Dumont E, et al. Monthly oral ibandronate is well tolerated and efficacious in postmenopausal women: results from the monthly oral pilot study. J Clin Endocrinol Metab. 2005 Sep;90(9):5018-24. doi: 10.1210/jc.2004-1750. PMID: 15972582.
121. Rhee CW, Lee J, Oh S, et al. Use of bisphosphonate and risk of atrial fibrillation in older women with osteoporosis. Osteoporos Int. 2012 Jan;23(1):247-54. doi: 10.1007/s00198-011-1608-z [doi]. PMID: 21431993.

© 2018 American Medical Association. All rights reserved.

122. Riis BJ, Ise J, von Stein T, et al. Ibandronate: a comparison of oral daily dosing versus intermittent dosing in postmenopausal osteoporosis. J Bone Miner Res. 2001 Oct;16(10):1871-8. doi: 10.1359/jbmr.2001.16.10.1871. PMID: 11585352.
123. Samelson EJ, Miller PD, Christiansen C, et al. RANKL inhibition with denosumab does not influence 3-year progression of aortic calcification or incidence of adverse cardiovascular events in postmenopausal women with osteoporosis and high cardiovascular risk. J Bone Miner Res. 2014;29(2):450-7.
124. Shiraki M, Fukunaga M, Kushida K, et al. A double-blind dose-ranging study of risedronate in Japanese patients with osteoporosis (a study by the Risedronate Late Phase II Research Group). Osteoporos Int. 2003 May;14(3):225-34. doi: 10.1007/s00198-002-1369-9. PMID: 12730746.
125. Simon JA, Recknor C, Moffett AH, Jr., et al. Impact of denosumab on the peripheral skeleton of postmenopausal women with osteoporosis: bone density, mass, and strength of the radius, and wrist fracture. Menopause. 2013 Feb;20(2):130-7. doi: 10.1097/gme.0b013e318267f909 [doi]. PMID: 23010883.
126. Sontag A, Wan X, Krege JH. Benefits and risks of raloxifene by vertebral fracture status. Curr Med Res Opin. 2010 Jan;26(1):71-6. doi: 10.1185/03007990903427082 [doi]. PMID: 19908937.
127. Sorensen HT, Christensen S, Mehnert F, et al. Use of bisphosphonates among women and risk of atrial fibrillation and flutter: population based case-control study. BMJ. 2008 Apr 12;336(7648):813-6. doi: 10.1136/bmj.39507.551644.BE. PMID: 18334527.
128. Tanko LB, Felsenberg D, Czerwinski E, et al. Oral weekly ibandronate prevents bone loss in postmenopausal women. J Intern Med. 2003 Aug;254(2):159-67. PMID: 12859697.
129. Thiebaud D, Burckhardt P, Kriegbaum H, et al. Three monthly intravenous injections of ibandronate in the treatment of postmenopausal osteoporosis. Am J Med. 1997 Oct;103(4):298-307. PMID: 9382122.
130. Tucci JR, Tonino RP, Emkey RD, et al. Effect of three years of oral alendronate treatment in postmenopausal women with osteoporosis. Am J Med. 1996 Nov;101(5):488-501. PMID: 8948272.
131. van Staa T, Abenhaim L, Cooper C. Upper gastrointestinal adverse events and cyclical etidronate. Am J Med. 1997 Dec;103(6):462-7. PMID: 9428828.
132. Vestergaard P, Schwartz K, Pinholt EM, et al. Gastric and esophagus events before and during treatment of osteoporosis. Calcif Tissue Int. 2010 Feb;86(2):110-5. doi: 10.1007/s00223-009-9323-x [doi]. PMID: 19957165.
133. Vestergaard P, Schwartz K, Pinholt EM, et al. Stroke in relation to use of raloxifene and other drugs against osteoporosis. Osteoporos Int. 2011 Apr;22(4):1037-45. doi: 10.1007/s00198-010-1276-4 [doi]. PMID: 20449570.
134. Vestergaard P. Acute myocardial infarction and atherosclerosis of the coronary arteries in patients treated with drugs against osteoporosis: calcium in the vessels and not the bones? Calcif Tissue Int. 2012 Jan;90(1):22-9. doi: 10.1007/s00223-011-9549-2 [doi]. PMID: 22120197.
135. Vestergaard P, Schwartz F, Rejnmark L, et al. Risk of femoral shaft and subtrochanteric fractures among users of bisphosphonates and raloxifene. Osteoporos Int. 2011 Mar;22(3):993-1001. doi: 10.1007/s00198-010-1512-y [doi]. PMID: 21165600.

© 2018 American Medical Association. All rights reserved.

136. Vestergaard P, Schwartz K, Rejnmark L, et al. Oral bisphosphonate use increases the risk for inflammatory jaw disease: a cohort study. J Oral Maxillofac Surg. 2012 Apr;70(4):821-9. doi: 10.1016/j.joms.2011.02.093. PMID: 21764202.
137. Rud B, Jensen JE, Mosekilde L, et al. Performance of four clinical screening tools to select peri- and early postmenopausal women for dual X-ray absorptiometry. Osteoporos Int. 2005 Jul;16(7):764-72. doi: 10.1007/s00198-004-1748-5. PMID: 15986263.

© 2018 American Medical Association. All rights reserved.