Osteoporosis International
https://doi.org/10.1007/s00198-019-04919-6

REVIEW ARTICLE

Performance of predictive tools to identify individuals at risk of non-traumatic fracture: a systematic review, meta-analysis, and meta-regression

C. Beaudoin[1,2,3] · L. Moore[1,2] · M. Gagné[3] · L. Bessette[2,4] · L. G. Ste-Marie[5] · J. P. Brown[2,4] · S. Jean[3,4]

Received: 18 October 2018 / Accepted: 26 February 2019

Abstract

Summary There is no consensus on which tool is the most accurate to assess fracture risk. The results of this systematic review suggest that QFracture, Fracture Risk Assessment Tool (FRAX) with BMD, and Garvan with BMD are the tools with the best discriminative ability. More studies assessing the comparative performance of current tools are needed.

Introduction Many tools exist to assess fracture risk. This review aims to determine which tools have the best predictive accuracy to identify individuals at high risk of non-traumatic fracture.

Methods Studies assessing the accuracy of tools for prediction of fracture were searched in MEDLINE, EMBASE, Evidence-Based Medicine Reviews, and Global Health. Studies were eligible if discrimination was assessed in a population independent of the derivation cohort. Meta-analyses and meta-regressions were performed on areas under the ROC curve (AUCs). Gender, mean age, age range, and study quality were used as adjustment variables.

Results We identified 53 validation studies assessing the discriminative ability of 14 tools. Given the small number of studies on some tools, only FRAX, Garvan, and QFracture were compared using meta-regression models. In the unadjusted analyses, QFracture had the best discriminative ability to predict hip fracture (AUC = 0.88). In the adjusted analysis, FRAX with BMD (AUC = 0.81) and Garvan with BMD (AUC = 0.79) had the highest AUCs. For prediction of major osteoporotic fracture, QFracture had the best discriminative ability (AUC = 0.77). For prediction of osteoporotic or any fracture, FRAX with BMD and Garvan with BMD had higher discriminative ability than their versions without BMD (FRAX: AUC = 0.72 vs 0.69, Garvan: AUC = 0.72 vs 0.65). A significant amount of heterogeneity was present in the analyses.

Conclusions QFracture, FRAX with BMD, and Garvan with BMD have the highest discriminative performance for predicting fracture. Additional studies in which the performance of current tools is assessed in the same individuals may be performed to confirm this conclusion.

Keywords Discrimination · Fracture · Osteoporosis · Risk assessment · Systematic review · Validation

Introduction

Osteoporosis is a systemic skeletal disease characterized by low bone mineral density (BMD) and deterioration of bone architecture resulting in reduced bone strength and, consequently, an increased susceptibility to fractures [1]. According to Osteoporosis Canada, at least one third of women and one fifth of men will suffer a fracture from osteoporosis during their

✉ C. Beaudoin
claudia.beaudoin@crchudequebec.ulaval.ca

[1] Department of Social and Preventive Medicine, Medicine Faculty, Laval University, Ferdinand Vandry Pavillon, 1050 Avenue de la Médecine, Quebec City, QC G1V 0A6, Canada

[2] CHU de Québec-Université Laval Research Center, Québec, QC, Canada

[3] Bureau d'information et d'études en santé des populations, Institut National de Santé Publique du Québec, 945, Avenue Wolfe, Québec G1V 5B3, Canada

[4] Department of Medicine, Medicine Faculty, Laval University, Ferdinand Vandry Pavillon, 1050 Avenue de la Médecine, Quebec City, QC G1V 0A6, Canada

[5] Department of Medicine, Medicine Faculty, University of Montréal, Montréal, QC, Canada

Published online: 14 March 2019

lifetime [2]. As fractures are associated with a decrease in quality of life [3, 4] and an increase in premature mortality [5–7], disability, and financial burden [8], it is important to identify individuals at high risk of fracture and to provide them with adequate therapeutic options.

In patients who have not yet suffered a fracture, measurement of BMD by dual-energy X-ray absorptiometry (DXA) is commonly used to identify those with osteoporosis or low BMD. However, many studies have shown that BMD measurement alone does not reliably predict if an individual will experience a fracture [9]. Therefore, predictive tools integrating numerous risk factors are now recommended in many clinical guidelines for identifying individuals at high risk of osteoporotic fracture [10–12].

Some reviews have been performed to assess if the performance of available tools is sufficient for practical use [13–16]. However, most of these suffer from a number of important limitations such as restriction to specific tools [16], restriction to cohorts of women [14], or assessment of calibration but not discrimination [15]. In the single study which includes meta-analyses [13], QFracture was the tool with the highest accuracy. However, a high level of heterogeneity was observed, which indicates that the difference in performance may be attributable to difference in study characteristics. Meta-regression models could have been used to adjust for the study characteristics and increase comparability between the tools, but the authors did not consider this approach. In addition, the only review including a meta-analysis was updated in September 2014 and therefore excluded several important validation studies published in the last few years [17–30]. Given the low number of studies on some tools, the inclusion of the most recent validation studies may provide a more accurate estimation of their performance.

In order to address the limitations of previous reviews and update available evidence, we performed a systematic review of the literature to identify all studies assessing the external validity of a fracture risk assessment tool. The objective of this review is to determine which tools have the best predictive accuracy to identify individuals at high risk of non-traumatic fracture among adults from the general population.

Methods

This systematic review was developed according to recommendations from the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy [31]. Eligibility criteria, outcomes, data to be extracted, quality assessment, and statistical analyses were determined a priori and formulated in an unpublished protocol. The results are presented according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [32].

Eligibility criteria

Prospective and retrospective cohort studies evaluating or comparing the accuracy of tools for prediction of fracture were included if predictive accuracy was assessed in a population independent of the derivation cohort. Internal validation studies were excluded since the performance of a model tends to be overestimated when evaluated using derivation data rather than new data [33]. Studies were eligible if the area under the ROC curve (AUC) and sufficient information to estimate its standard error were reported [34]. The validation cohort had to include adults from the general population. Tools may have been constructed to predict an outcome other than fracture (e.g., low BMD) but should have been validated for fracture prediction. Cross-sectional and case-control studies as well as validation studies in which all the participants had a given disease (e.g., osteoporosis, cancer) or were on specific medication (e.g., corticosteroids) were excluded. Studies assessing the ability of a tool to predict pathological fracture caused by a disease other than osteoporosis were also excluded. Studies published in any language, peer-reviewed or not, were included.

Search strategy

The search strategy is based on the strategy of Nayak et al. [15] and was revised by a professional research librarian (Vicky Tessier, Institut National de Santé Publique du Québec). The electronic databases MEDLINE, EMBASE, Evidence-Based Medicine Reviews, and Global Health were searched via OvidSP from their inception through August 2017 for validation studies of fracture risk assessment tools. The search strategy was designed for MEDLINE (Appendix Table 9) and then adapted for the other databases using a combination of MeSH, EMTREE, or CAB Thesaurus terms and keywords related to the terms "fracture," "osteoporosis," and "risk prediction tool." No filter for diagnostic studies was used in order to maximize the sensitivity of the search. References from identified articles and existing reviews were screened. All references were imported into EndNote software (version X7.0.2, Thomson Reuters), and duplicates were removed.

Study selection

Based on the inclusion and exclusion criteria described above, three reviewers (CB, SJ, and MG) independently screened titles and abstracts to determine study eligibility. Each article was screened by two different reviewers. The full text of studies classified as eligible or unclear were also independently assessed for final inclusion by two reviewers (CB and SJ). Disagreements were discussed and, when required, resolved by a forth reviewer (LM).

Data abstraction

Data from included studies were extracted separately by two reviewers (CB and SJ) using previously standardized data extraction forms piloted on five representative studies. Extracted information included derivation and validation cohorts (i.e., country, setting, follow-up period, number of participants), participants' clinical characteristics (i.e., age, gender, treatment for osteoporosis, history of fracture), tool description (i.e., name, risk factors, osteoporotic fracture definition, length of the prediction period) as well as performance measures (i.e., AUC and their standard errors). A formula from Hanley and McNeil [34] was used to approximate unreported standard errors of AUCs. Some authors were contacted in order to obtain additional data (i.e., AUCs and their standard errors).

Quality assessment

Two reviewers (CB and SJ) independently assessed the quality of selected studies using a checklist derived from the Quality Assessment of Diagnostic Accuracy Study 2 (QUADAS-2) tool [35], adapted by the project steering committee (Appendix 2). Disagreements were discussed and, when required, resolved by the forth reviewer (LM). We considered that a potential selection bias exists if less than 90% of eligible patients were included in the analyses [36]. We also used a rule of thumb of a minimum of 10 events per predictor to establish sufficient sample size [37, 38].

Data synthesis and analyses

Meta-analyses

To summarize the discriminative ability of each tool from multiple studies, a summary AUC value was calculated under a random-effects model using the metafor package of the R software [39]. An AUC between 0.7 and 0.8 was considered acceptable, between 0.8 and 0.9 excellent, and higher than 0.9 outstanding [40]. Pooled and individual measures and their 95% confidence intervals (95% CI) were illustrated in forest plots. Results were presented by length of prediction period and osteoporotic fracture definition. Heterogeneity was quantified using the I^2 statistic and the thresholds proposed in the Cochrane Handbook for Systematic Reviews of Interventions (0–40%: low, 30–60%: moderate, 50–90%: substantial, 75–100%: considerable heterogeneity) [41].

Meta-regressions

Meta-regression models were also conducted with the metafor package, using the AUC as a dependent variable and the tool as independent variable. With this method, comparison of the discriminative ability of each tool was performed by considering the correlation between results from the same study. In order to improve the comparability of different studies evaluating different tools, some key characteristics of the validation cohort determined a priori (gender, mean age, age range, and global study quality) were used as covariables. Age range was defined as the maximal age value minus the minimal age value of individuals within a study. In order to achieve model convergence, only the tools for which more than two AUCs were available were included in the meta-regressions. Residual heterogeneity was tested.

Results

Among the 32,167 studies identified using our search strategy, 53 satisfied the eligibility criteria and were included in the meta-analyses (Fig. 1). Only 44 studies on tools externally validated more than twice (FRAX, Garvan, QFracture) could be included in the meta-regression models. The 53 eligible studies included more than seven million participants from 20 countries in North America, Europe, Africa, Asia and, Australia (Table 1). Among these studies, only 5 included exclusively men, 18 included both men and women, and 30 included exclusively women. Mean age varied from 47 to 79 years, age range from 12 (69–81 years in study [18]) to 69 years (30–99 years in study [71]), and mean length of follow-up from 3 to 11 years.

Among the 53 included studies, 19 were of low quality for at least three domains of the QUADAS-2 tool and were thus attributed a low overall quality (Appendix Fig. 2). In 17 studies, more than 10% of data was missing and imputed to a null or average value. More than 10% of eligible individuals were excluded from the analysis in 27 studies. Fractures were not confirmed using an additional source of data (e.g., radiological reports, surgical reports, medical records) in 46% of the studies. Half of the studies had insufficient sample size with less than ten events per predictor for at least one of the outcomes. For 36% of the studies, the follow-up time was unknown or different from the length of the period for which the tool was intended.

In the 53 eligible studies, 14 different tools were validated: the Fracture Risk Assessment Tool (FRAX), the Garvan Fracture Risk Calculator (Garvan), the original version of QFracture (QFracture 2009), the updated version of QFracture (QFracture 2012), the Fracture Risk Calculator (FRC), the Women Health Initiative (WHI) hip fracture risk score, the Osteoporosis Index of Risk (OSIRIS), the Osteoporosis Risk Assessment Instrument (ORAI), the Osteoporosis Self-Assessment Tool (OST), the Simple Calculated Osteoporosis Risk (SCORE), the Fracture and Immobilization Score (FRISC), the modified version of the Study of Osteoporotic Fracture (mSOF) frailty index, the FRActure Health Search (FRA-HS) score, and the Fracture

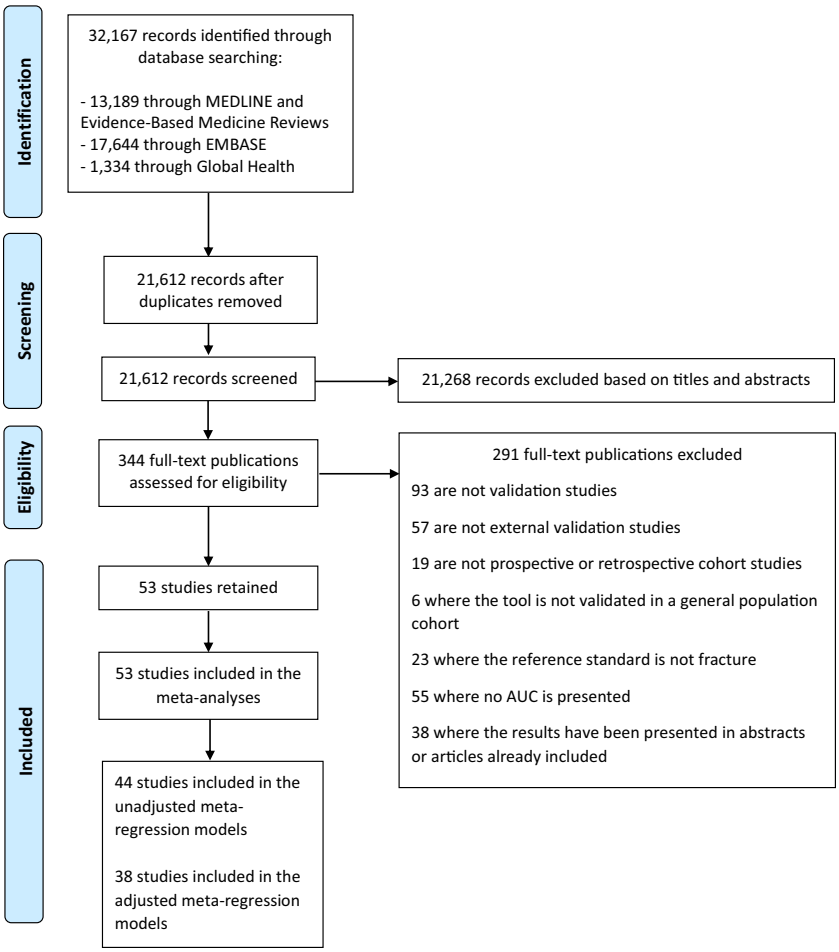


Fig. 1 Flow diagram of studies

and Mortality (FRAMO) index. The tools included between 2 (OST) and 30 (QFracture 2012) clinical risk factors (Table 2). The most common risk factors included in prediction models were age, weight, and history of fracture. Nine tools were originally developed to estimate fracture risk (FRAX, Garvan, QFracture, FRC, WHI score, FRISC, FRA-HS, FRAMO); four tools were developed for identification of patients with low BMD (OSIRIS, ORAI, OST, and SCORE); one estimated risk of death (FRAMO), and one was conceived to assess frailty (mSOF). The risk of fracture was estimated on prediction periods between 1 and 10 years.

Performance of tools to predict hip fracture

In the 37 studies included in the meta-analyses on the performance of tools for prediction of hip fracture (Table 3), AUCs were available for FRAX, Garvan, QFracture 2009, QFracture 2012, FRC, WHI score, FRA-HS, and FRAMO. The discriminative ability of the tools was acceptable or excellent, with AUCs varying between 0.70 (Garvan without BMD) and 0.88 (QFracture 2009). Heterogeneity between AUCs was considerable for most tools. Only FRAX, Garvan, and QFracture could be compared through meta-regression models (Table 4). In the unadjusted model, Garvan without BMD had the lowest discriminative ability (AUC [95% CI] = 0.72 [0.66, 0.78] for 10-year prediction) and QFracture 2009 had the highest (AUC [95% CI] = 0.88 [0.86, 0.89]). After adjustment for the study characteristics, Garvan without BMD still had the lowest discriminative ability (AUC [95% CI] = 0.70 [0.64, 0.76] for 10-year prediction), but QFracture 2009 was not superior to the other tools. AUCs were significantly higher when BMD was used in combination with other risk factors to predict hip fracture. AUCs were significantly lower in studies with higher mean age (mean decrease [95% CI] in AUC of 0.0043 [0.0017, 0.0068] for every 1-year increase in the mean age, $p = 0.0010$) and lower age range (mean increase [95% CI] in AUC of 0.0026 [0.0007, 0.0045] for every 1-year increase in the age range, $p = 0.0076$). AUCs tend to be lower in cohorts of men than women (mean difference [95% CI] of 0.0122 [− 0.0200, 0.0445] between both sexes, $p = 0.46$), but the difference did not reach statistical significance. No significant association was detected between AUC and study quality. A significant amount of residual heterogenity was present in the adjusted model ($p < 0.0001$).

Table 1 Characteristics of studies assessing the external validity of fracture risk assessment tools

Studies	Countries	Validation cohort	Years of follow-up	Type of study	Mean follow-up (min–max) (years)	Number of participants	Sex	Mean age (min–max) (years)
Pressman 2011 [42]	USA	Members of Kaiser Permanente Northern California	1997–NA	R	6.6 (1.0–10.0)	94,489	W	62.0* (50–85)
Lo 2011 [43]			1997–2009	R	6.6 (1.0–10.0)	94,489	W	62.8 (50–85)
Ettinger 2012 [44]	USA	Mr. OS and/or Ms. OS Studies	2000–2012	P	8.4 (7.3–10.0)	5893	M	73.6 (65–NA)
Ettinger 2013 [45]	USA		2000–2012	P	8.4 (7.3–10.0)	5891	M	73.6 (65–NA)
Su 2017 [30]	China		2001–NA	P	9.9 (NA), 8.8 (NA)	1923, 1950	M, W	72.3 (65–NA), 72.5 (65–NA)
Orwoll 2017 [28]	USA, Sweden, China		NA	P	8.6 (NA), 10.6 (NA), 9.8 (NA)	1469, 2542, 1476	M	75.0 (65–NA), 74.0 (65–NA), 72.0 (65–NA)
Yu 2014 [46]	China		2001–2013	P	10.2 (NA)	2000, 2000	M, W	72.4 (65–NA), 72.6 (65–NA)
Cordomí 2013 [47]	Spain	Centre for Technical Studies with Radioactive Isotopes database	1992–NA	R	11.0* (NA)	1231	W	56.8 (40–90)
Czerwiński 2013 [48]	Poland	Women who had a densitometric examination in Cracow Medical Centre	1997–NA	P	11.0 (NA)	1024	W	63.8 (50–79)
Sornay-Rendu 2010 [49]	France	OFELY cohort	1992–2003	P	NA (NA–10.0)	867	W	58.8 (40–89)
Cheung 2012 [50]	China	Hong Kong Osteoporosis Study	1995–NA	P	4.5 (1.0–14.6)	2266	W	62.1 (41–90)
Trémollieres 2010 [51]	France	Menopause et Os cohort study	1988–2004	P	NA	956	W	NA (45–NA)
Yun 2010 [52]	USA	Medicare Current Beneficiary Survey	1999–2005	R	NA (NA–2.0)	12,413	W, M	NA (65–NA)
González-Macías 2012 [53]	Spain	Women recruited in 58 primary care centers of the National Health Service	2000–NA	P	3.0* (NA)	5120	W	72.3 (65–100)
Tamaki 2011 [54]	Japan	Japanese Population-Based Osteoporosis Cohort Study	1996–2006	P	NA (NA–10)	815	W	56.7 (40–74)
Briot 2013 [55]	UK, Germany, France	Osteoporosis and Ultrasound Study	1999–NA	P	6.0 (4.5–7.5)	1748	W	66.1 (55–NA)
Ensrud 2009 [56]	USA	Study of Osteoporotic Fractures	1986–NA	P	NA	6252	W	71.3 (65–NA)
Rubin 2013 [57]	Denmark	Danish National Civil Registration System	2009–2012	P	3.0 (2.5–3.1)	3614	W	64.0 (40–90)
Langsetmo 2011 [58]	Canada	Canadian Multicentre Osteoporosis Study	NA–2006	P	8.6 (1.0–10.0), 8.3 (1.0–10.0)	4152, 1606	W, M	67.7 (55–95), 67.6 (55–95)
Bolland 2011 [59]	New Zealand	Placebo-controlled trial of calcium supplements	1998–2009	P	8.8 (0.2–11.4)	1422	W	74.2 (55–NA)
Azagra 2016 [19]	Spain	FRIDEX cohort	2000–2010	P	NA	770	W	56.8 (40–90)
Henry 2011 [60]	Australia	Geelong Osteoporosis Study	1994–NA	P	9.6* (NA)	600	W	74.0 (60–NA)
Zhang 2011 [61]			1993–NA	P	9.2* (NA)	587	W	NA (60–90)
Holloway 2016 [21]			2001–2015	P	9.5* (NA)	591	M	70.0 (40–90)

Table 1 (continued)

Studies	Countries	Validation cohort	Years of follow-up	Type of study	Mean follow-up (min–max) (years)	Number of participants	Sex	Mean age (min–max) (years)
Leslie 2010 [62]	Canada	Manitoba Bone Density Program database	1990–2008	R	NA	39,603	W, M	65.9 (50–NA)
Morin 2009 [63]			1998–2004	R	3.3 (NA)	8252	W	52.7 (40–59)
Hundrup 2010 [64]	Denmark	Danish Nurse Cohort	1999–2004	P	NA (NA–5.0)	13,353	W	61.0 (50–79)
Hundrup 2009 [65]			1999–2004	P	NA (NA–5.0)	13,353	W	61.0 (50–79)
Lundin 2015 [18]	Sweden	PRIMary health care and Osteoporosis project	1999–NA	P	9.9 (NA)	388	W	73.2 (69–81)
Albaba 2014 [66]	USA	Adults enrolled in a primary care panel in Olmsted County	2005–2008	R	NA (NA–4.0)	850	W, M	NA (60–NA)
Cauley 2010 [67]	USA	Clinical Trials or Observational Study of the Women's Health Initiative	NA	P	10.3 (NA)	159,579	W	NA (NA)
Robbins 2007 [68]			NA–2005	P	8.0 (NA)	68,132	W	62.7 (50–79)
Ahmed 2014 [69]	Norway	Tromsø Study	2001–2009	P	6.9 (NA), 7.1 (NA)	1637, 1355	W, M	69.3 (60–84), 69.7 (60–84)
Tanaka 2010 [70]	Japan	Miyama and Taiji Cohorts	1990–2003	P	NA	400	W	59.5 (40–79)
Hippisley-Cox 2014 [71]	UK	Clinical Practice Research Datalink	1998–2012	R	NA (NA–14.6)	2,852,381	W, M	NA (30–99)
Klop 2016 [22]			2004–2013	R	NA (NA–10.0)	24,227	W, M	63.0 (40–90)
Collins 2011 [72]	UK	THIN database	1994–2008	R	6.0 (NA–14)	1,136,417, 1,108,219	W, M	48.0 (30–85), 47.0 (30–85)
Hippisley-Cox 2009 [73]	England and Wales	Patients registered from 178 general practices that use the Egton Medical Information System computer system	1993–2008	R	NA (NA)	454,499, 424,336	W, M	NA (40–85), NA (40–85)
van Geel 2014 [74]	Netherlands	10 general practice centers located in the southern part of The Netherlands	1992–2004	P	NA (NA)	506	W	67.8 (60–80)
Friis-Holmberg 2014 [75]	Denmark	Danish Health Examination Survey 2007–2008	2007–2012	P	4.3 (0.03–4.9)	7552, 5206	W, M	56.8 (40–90)58.3 (40–90)
Forti 2012 [76]	Italy	Conselice Study of Brain Aging	1999–2004	P	3.8 (NA)	714	W, M	74.7 (65–NA)
Tanaka 2011 [77]	Japan	Nagano Cohort	1993–2011	P	5.1 (NA)	765	W	63.3 (45–81)
Sambrook 2011 [78]	Australia, Belgium, Canada, France, Germany, Italy, Netherlands, Spain, UK, USA	Global Longitudinal Study of Osteoporosis in Women	2006–NA	P	NA (NA–2.0)	19,586	W	NA (60–NA)
Dagan 2017 [24]	Israel	Members of Clalit Health Services	2010–2014	R	4.7 (NA–5.0)	1,054,815	W, M	65.8 (50–90)
Dominguez 2017 [29]	Spain	Men and women from a Canary Island population	NA	P	NA (NA–10.0)	121	W, M	59.3 (NA)
Olmos 2017 [27]	Spain	Cohort of Spanish postmenopathic women	NA	P	6.8 (NA)	1809	W	64.0 (44–90)
Hoff 2017 [26]	Norway	Nord-Trondelag Health study	2006–NA	P	5.2* (NA)		W, M	

Table 1 (continued)

Studies	Countries	Validation cohort	Years of follow-up	Type of study	Mean follow-up (min–max) (years)	Number of participants	Sex	Mean age (min–max) (years)
						15,432, 13,585		64.4 (50–90), 64.0 (50–90)
Francesco 2017 [25]	Italy	"Mille in Rete" database	2003–2015	R	NA (NA–10.0)	82,242	W, M	57.7 (40–NA)
Sundh 2017 [23]	Sweden	Prospective Population Study of Women in Gothenburg	1980–1993	P	10.0 (NA)	499	W	55.1 (50–66)
Goldshtein 2016 [20]	Israel	Patients insured under a large Israeli health fund	NA	R	NA (NA–10.0)	6000	W, M	61.0 (50–NA)
Iki 2015 [17]	Japan	Fujiwara-kyo Osteoporosis Risk in Men Cohort Study	2007–2014	P	4.5* (NA)	1805	M	73.0 (65–NA)
Sund 2014 [79]	Finland	Kuopio Osteoporosis Risk Factor and Prevention Study	1994–2004	P	NA (NA)	2755	W	59.1 (53–66)
Albertsson 2007 [80]	Sweden	National Swedish Population Register	2002–2003	R	NA (NA–2.0)	1248	W	78.8 (70–100)

*Median is presented instead of mean

P prospective, *R* retrospective, *W* women, *M* men, *NA* unavailable data

Performance of tools to predict FRAX-defined major osteoporotic fracture or QFracture-defined osteoporotic fracture

In the 34 studies included in the meta-analysis on the performance of tools to predict FRAX-defined major osteoporotic fracture (hip, clinical spine, wrist, or humerus) or QFracture 2009–defined osteoporotic fracture (hip, vertebral, or wrist) or QFracture 2012–defined osteoporotic fracture (hip, vertebral, humerus, or wrist) (Table 5), AUC was presented for 11 tools (FRAX, Garvan, QFracture 2009, QFracture 2012, FRC, OSIRIS, ORAI, OST, SCORE, FRISC, FRA-HS). The discriminative ability of the tools was low to acceptable (AUC between 0.63 and 0.79). Heterogeneity in AUC values was again substantial for most tools (I^2 between 71.5 and 99.7%). In both unadjusted and adjusted meta-regressions (Table 6), QFracture 2009 was the tool with the best discriminative ability. The version of FRAX including BMD had significantly greater discriminative ability than the version without BMD. Performance was significatively lower in validation studies including men exclusively (mean difference [95% CI] in AUC of 0.0445 [0.0154, 0.0735] between both sexes, $p = 0.0027$). No significant association was detected between AUC and mean age, age range, or study quality. Meta-regression model accounting for gender, mean age, age range, and study quality did not completely explain heterogeneity ($p < 0.0001$).

Performance of tools to predict Garvan-defined osteoporotic fracture, fragility fracture, or any fracture

Only 11 studies were included in the meta-analyses on the performance of tools to predict Garvan-defined osteoporotic fracture (hip, clinical spine, wrist, humerus, meta-carpal, scapula, clavicle, distal femur, tibia/fibula, patella, pelvis, or sternum), fragility fracture, or any fracture. These studies evaluated the discriminative ability of FRAX, Garvan, OSIRIS, OST, ORAI, and SCORE. The discriminative ability was low (Table 7), with AUCs between 0.60 (FRAX without BMD) and 0.69 (ORAI and SCORE). In both unadjusted and adjusted meta-regression models (Table 8), discriminative ability was significantly higher for FRAX and Garvan with BMD than for FRAX and Garvan without BMD. AUCs significantly decreased with mean age of the validation cohort (mean decrease [95% CI] in AUC of 0.0085 [0.0008, 0.0161] for every 1-year increase in the mean age, $p = 0.03$), but did not change significantly with age range, gender, or study quality. A significant amount of heterogeneity remained after adjustment for the characteristics of the studies ($p = 0.0033$).

Table 2 Description of the externally validated tools for prediction of fracture

	FRAX	Garvan	QFracture2009	QFracture2012	FRC	WHI score	OSIRIS	ORAI	OST	SCORE	FRISC	mSOF index	FRA-HS score	FRAMO index
Derivation cohort														
Sex of participants	M, W	M, W	M, W	M, W	M, W	W	W	W	W	W	W	M, W	M, W	W
Age range (years)	40–106	60+	30–85	30–100	45+	50–79	60–80	45+	45–88	45+	PM	65+	40+	70–100
Clinical risk factor (predictor)														
Age	X	X	X	X	X	X	X	X	X	X	X		X	X
Sex	X	X	X	X	X								X	
Weight	X	(X)	X	X	X	X	X	X	X	X	X		X	X
Height	X		X	X	X	X							X	
History of fx	X	X		X	X	X	X			X	X		X	X
Parental history of OP or hip fx	X		X	X	X	X								
Tobacco smoking	X		X	X	X	X							X	
Use of glucocorticoids	X		X	X	X	X							X	
Rheumatoid arthritis	X		X	X	X					X			X	
Secondary osteoporosis	X				X								X	
Alcohol consumption	X		X	X	X								X	
BMD	(X)	(X)			(X)						X			
History of falls		X	X	X										
Ethnicity				X		X				X				
Global health						X								
Physical activity						X								
Impaired raise up												X		X
Energy level												X		
Malnutrition												X		
Menopausal symptoms			X	X										
Diabetes			X	X		X								
Back pain											X			
Dementia				X										
Cardiovascular disease			X	X										
Asthma			X	X										
COPD				X										
Chronic liver disease			X	X										
Gastrointestinal malabsorption			X	X										
Endocrine problems			X	X										
Parkinson's disease				X										

Table 2 (continued)

	FRAX	Garvan	QFracture2009	QFracture2012	FRC	WHI score	OSIRIS	ORAI	OST	SCORE	FRISC	mSOF index	FRA-HS score	FRAMO index
Cancer				X										
Chronic renal disease				X										
Systemic lupus erythematosus				X										
Epilepsy				X										
Living in a care or nursing home				X										
Hormone replacement therapy			X	X			X	X		X				
Antidepressive drugs			X	X										
Anticonvulsive drugs				X										
Outcome														
Hip fx	X	X	X	X	X	X							X	X
Major osteoporotic fx	X		X	X	X						X		X	
Any fx		X												
BMD							X	X	X	X				
Frailty												X		
Death														X
Length of the prediction period (years)	10	5 or 10	1–10	1–10	5	5	NA	NA	NA	NA	1, 3, 5 or 10	NA	10	2

FRAX Fracture Risk Assessment Tool, *Garvan* Garvan Fracture Risk Calculator, *FRC* Fracture Risk Calculator, *WHI* Women Health Initiative, *OSIRIS* Osteoporosis Index of Risk, *ORAI* Osteoporosis Risk Assessment Instrument, *OST* Osteoporosis Self-Assessment Tool, *SCORE* Simple Calculated Osteoporosis Risk, *FRISC* Fracture and Immobilization Score, *mSOF index* modified Study of Osteoporotic Fracture frailty index, *FRA-HS* FRActure Health Search, *FRAMO* Fracture and Mortality, *M* men, *W* women, *PM* postmenopausal, *OP* osteoporosis, *fx* fracture, *COPD* chronic obstructive pulmonary disease, *(X)* optional, *NA* not applicable (no follow-up when BMD or frailty is the outcome)

Table 3 Summary AUCs for prediction of hip fracture, results of meta-analyses. *AUC* area under the ROC curve, *CI* confidence interval, I^2 measure of heterogeneity

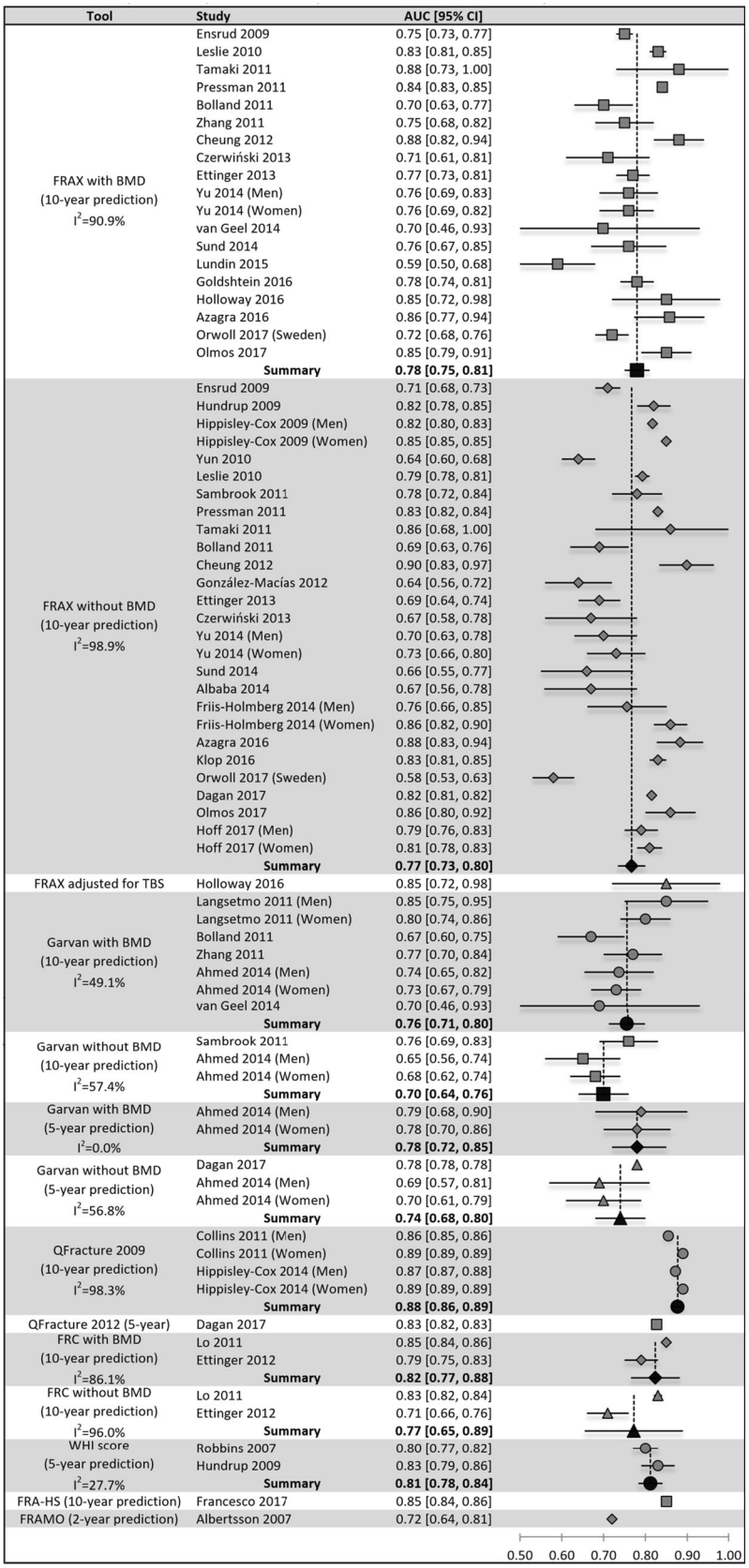


Table 4 Summary AUC for prediction of hip fracture, results of meta-regression models

Tools	Unadjusted model[g]		Adjusted model[g,h]	
	n	AUC [95% CI]	*n*	AUC [95% CI]
FRAX with BMD (10-year prediction)	19	0.79 [0.77, 0.81][b,d,e,f]	19	0.81 [0.78, 0.84][b,d,e]
FRAX without BMD (10-year prediction)	27	0.75 [0.72, 0.79][a,e,f]	24	0.77 [0.73, 0.81][a,d,e]
Garvan with BMD (10-year prediction)	7	0.77 [0.73, 0.81][f]	7	0.79 [0.74, 0.84][d]
Garvan without BMD (10-year prediction)	3	0.72 [0.66, 0.78][a,f]	2	0.70 [0.64, 0.76][a,b,c]
Garvan without BMD (5-year prediction)	3	0.74 [0.72, 0.76][a,b,f]	3	0.75 [0.72, 0.79][a,b]
QFracture 2009 (10-year prediction)	4	0.88 [0.86, 0.89][a,b,c,d,e]	4	0.75 [0.68, 0.81]

AUC area under the ROC curve, *CI* confidence interval, *n* number of AUCs

[a] Statistically different from AUC for FRAX with BMD (10-year prediction)

[b] Statistically different from AUC for FRAX without BMD (10-year prediction)

[c] Statistically different from AUC for Garvan with BMD (10-year prediction)

[d] Statistically different from AUC for Garvan without BMD (10-year prediction)

[e] Statistically different from AUC for Garvan without BMD (5-year prediction)

[f] Statistically different from AUC for QFracture 2009 (10-year prediction)

[g] Test for residual heterogeneity: $p < 0.0001$

[h] Meta-regression model adjusted for gender, mean age, age range, and global quality. AUCs are presented for studies of high quality (score = 9/9) and cohort of women whose mean age is 65 years and age range is 40 years

Discussion

In this systematic review, we identified 53 external validation studies assessing the discriminative ability of 14 different tools for fracture prediction (FRAX, Garvan, QFracture 2009, QFracture 2012, FRC, WHI hip fracture risk score, OSIRIS, ORAI, OST, SCORE, FRISC, mSOF index, FRA-HS score, and FRAMO index). Overall quality was considered as low in about 35% of studies. Sufficient data allowing comparisons using meta-regression models was available for FRAX, Garvan, and the original version of QFracture (QFracture 2009) only. Tools generally performed better at predicting hip fracture than major osteoporotic, osteoporotic, or any fracture. In the unadjusted analysis, QFracture 2009 had the highest discriminative ability to predict hip fracture, but in the adjusted analysis, FRAX with BMD (AUC = 0.81) and Garvan (AUC = 0.79) with BMD were the tools with the highest AUCs. QFracture 2009 (AUC = 0.77) had the highest discriminative ability to predict major osteoporotic fracture in both unadjusted and adjusted analyses. For prediction of osteoporotic fracture or any fracture, Garvan with BMD (AUC = 0.72) and FRAX with BMD (AUC = 0.72) had a higher discriminative ability than Garvan without BMD (AUC = 0.67) and FRAX without BMD (AUC = 0.69). AUCs tend to be higher in women and younger cohorts. For prediction of hip fracture, AUCs were significantly higher in cohorts with a broader age range.

Unadjusted analyses

In a recent meta-analysis by Marques et al. [13], QFracture had higher discriminative ability than FRAX and Garvan in predicting hip and major osteoporotic fracture. This result is consistent with the results of our meta-analyses and unadjusted meta-regression models. In the review by Marques et al. as well as in ours, a high level of heterogeneity was observed. Heterogeneity in AUCs for a same tool can be attributed to multiple factors such as level of homogeneity in the characteristics of individuals from the validation cohort (e.g., wide vs narrow age range, men or women only vs both sexes) or study quality (e.g., follow-up approximately the same as the tool's prediction period). Given that risk factors for fracture, their importance, and their prevalence may vary across subgroups [16, 81], predictive ability can also differ among subpopulations (e.g., younger vs older individuals, women vs men). An important limitation of the unadjusted analyses is that they compare different tools from different validation studies without consideration of the study characteristics affecting the predictive ability.

Adjusted meta-regression models

In our review, adjusted meta-regression models were also constructed in order to reduce heterogeneity and improve the comparability of the studies. To the best of our knowledge, we are the first to use these sophisticated models to compare the discriminative ability of current fracture risk prediction tools while adjusting for some key characteristics of the validation studies. Consistent with the results of Marques et al., the results of our adjusted meta-regression models suggest that QFracture is superior for prediction of major osteoporotic fracture. However, in contrast to the results of Marques et al., they do not offer evidence that QFracture is superior

Table 5 Summary AUC for prediction of FRAX-defined major osteoporotic fracture or QFracture-defined osteoporotic fracture, results of meta-analyses. FRAX-defined major osteoporotic fracture: hip, clinical spine, wrist, or humerus, QFracture-defined major osteoporotic fracture: hip, vertebral, or wrist. *AUC* area under the ROC curve, *CI* confidence interval, I^2 measure of heterogeneity

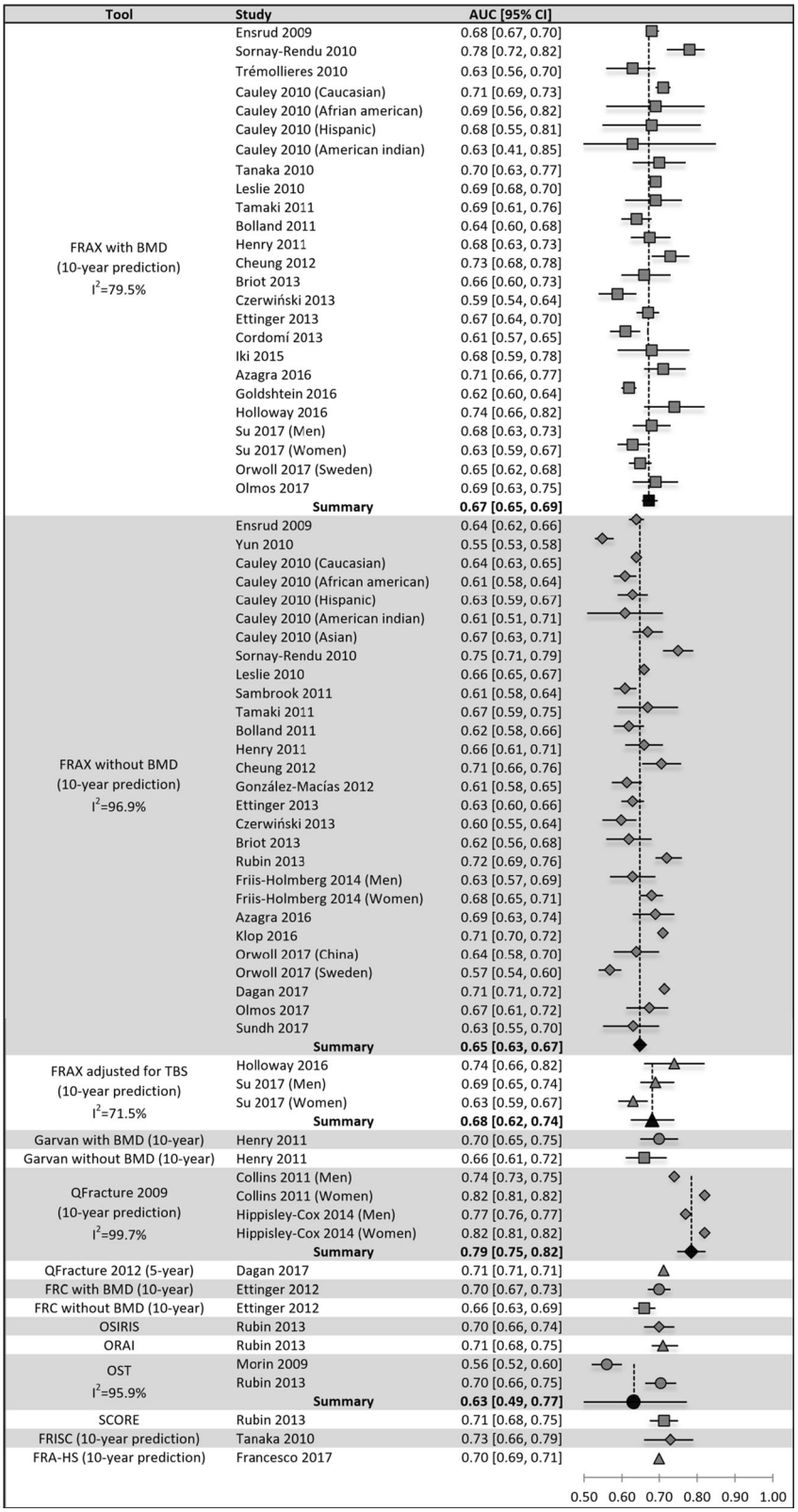


Table 6 Summary AUC for prediction of FRAX-defined major osteoporotic fracture or QFracture-defined osteoporotic fracture, results of meta-regression models

Tools	Unadjusted model[e]		Adjusted model[e,f]	
	n	AUC [95% CI]	*n*	AUC [95% CI]
FRAX with BMD (10-year prediction)	25	0.68 [0.66, 0.70][b,d]	20	0.70 [0.66, 0.74][b,d]
FRAX without BMD (10-year prediction)	28	0.64 [0.62, 0.66][a,c,d]	21	0.67 [0.63, 0.70][a,c,d]
FRAX adjusted for TBS (10-year prediction)	3	0.69 [0.65, 0.73][b,d]	3	0.71 [0.65, 0.77][b]
QFracture 2009 (10-year prediction)	4	0.79 [0.75, 0.82][a,b,c]	4	0.77 [0.70, 0.83][a,b]

FRAX-defined major osteoporotic fracture: hip, clinical spine, wrist, or humerus. QFracture 2009–defined major osteoporotic fracture: hip, vertebral, or wrist

AUC area under the ROC curve, *CI* confidence interval, *n* number of AUCs

[a] Statistically different from AUC for FRAX with BMD

[b] Statistically different from AUC for FRAX without BMD

[c] Statistically different from AUC for FRAX adjusted for TBS

[d] Statistically different from AUC for QFracture 2009

[e] Test for residual heterogeneity: $p < 0.0001$

[f] Meta-regression model adjusted for gender, mean age, age range, and global quality. AUCs are presented for studies of high quality (score = 9/9) and cohort of women whose mean age is 65 years and age range is 40 years

for prediction of hip fracture. This challenges the superiority of QFracture and raises the need to compare the predictive ability of current tools in the same group of individuals.

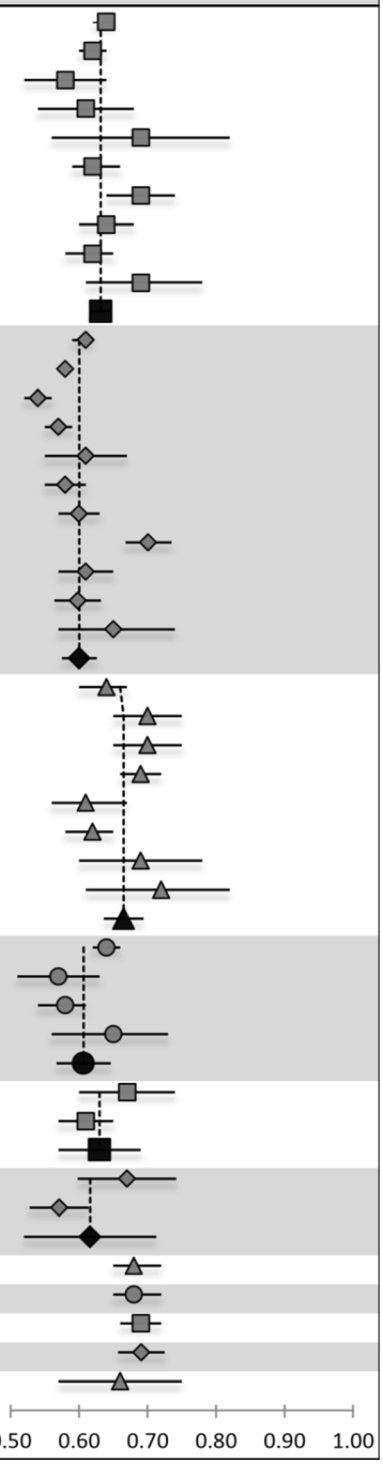
Consistent with the study of Marques et al., the results of our adjusted and unadjusted analyses also suggest that current tools perform better for prediction of hip fracture than major osteoporotic, osteoporotic, or any fracture. While it is clinically more convenient to interpret a single risk than one for each fracture site, the estimation of fracture risk at many sites simultaneously could decrease performance because risk factors for fracture, or the magnitude of their effect, differ across fracture sites [81, 82]. An additional result obtained through our adjusted analyses is that the discriminative ability of current tools tended to be higher in women and younger cohorts. The lower proportion of traumatic fracture and the better knowledge of the risk factors for fracture in women could explain why the discriminative ability is higher in this population. Studying risk factors for osteoprotic fracture in men remains important for improvement of fracture risk prediction in this population. A possible explanation for the higher discriminative ability observed in younger individuals is that the competing risk of death is higher in older than younger individuals, which translates into a higher proportion of older individuals at high risk of fracture who dies rather than suffers a fracture during the follow-up period. The development of a distinct model by age subgroups and the integration of new predictors such as frailty-related risk factors and polypharmacy could increase predictive ability in older populations.

In another systematic review by Rubin et al. [14], the authors concluded that tools including only a few risk factors have similar and sometimes higher predictive ability than more complex tools. The results of our review contrast with this conclusion and rather suggest that more complex tools tend to have higher discriminative ability than more simple tools. In their review, Rubin et al. made no distinction between versions with and without BMD. In our review, we observed that Garvan with BMD and FRAX with BMD had significantly higher discriminative ability than their version without BMD. Given the low number of external validation studies on the simplest tools (FRAMO, OST, ORAI, OSIRIS), it was difficult to compare them with other tools. There were also a low number of studies assessing the external validity of QFracture, the most complex tool. However, based on available evidence, the results of our review suggest that QFracture is among the best tools to predict fracture even if it does not use BMD. While BMD is a strong predictor of a future fracture, the inclusion of a large number of risk factors seems to compensate for the lack of this information. As the predictive ability tends to increase with the number of predictors, the inclusion of additional risk factors represents a potential avenue for improving the performance of fracture risk prediction tools.

The choice of the tool to use for prediction of fracture risk should be made by considering their performance in the targeted context, applicability, and characteristics. While the performance of QFracture was good, it was conceived to be used in electronic medical records and, to the best of our knowledge, has only been validated in this type of data. Furthermore, its original version is no longer available on the web and has been replaced by an updated version including 31 clinical risk factors. The updated QFracture may be cumbersome to use in a clinical context, unless it is integrated to electronic medical systems for automatic calculation of fracture risk. In a context where information on risk factors must be entered manually, FRAX and Garvan are much easier to use. The performance of these tools is good, especially for

Table 7 Summary AUC for prediction of Garvan-defined osteoporotic fracture, fragility fracture, or any fracture, results of meta-analyses. Garvan-defined osteoporotic fracture: hip, clinical spine, wrist, humerus, meta-carpal, scapula, clavicle, distal femur, tibia/fibula, patella, pelvis, and sternum. *AUC* area under the ROC curve, *CI* confidence interval, I^2 measure of heterogeneity

Tool	Study	AUC [95% CI]
FRAX with BMD (10-year prediction) I^2=29.8%	Ensrud 2009	0.64 [0.62, 0.65]
	Cauley 2010 (Caucasian)	0.62 [0.60, 0.64]
	Cauley 2010 (Afrian american)	0.58 [0.52, 0.64]
	Cauley 2010 (Hispanic)	0.61 [0.54, 0.68]
	Cauley 2010 (American indian)	0.69 [0.56, 0.82]
	Bolland 2011	0.62 [0.59, 0.66]
	Zhang 2011	0.69 [0.64, 0.74]
	Yu 2014 (Men)	0.64 [0.60, 0.68]
	Yu 2014 (Women)	0.62 [0.58, 0.65]
	van Geel 2014	0.69 [0.61, 0.78]
	Summary	**0.63 [0.62, 0.65]**
FRAX without BMD (10-year prediction) I^2=93.2%	Ensrud 2009	0.61 [0.59, 0.62]
	Cauley 2010 (Caucasian)	0.58 [0.58, 0.58]
	Cauley 2010 (African american)	0.54 [0.52, 0.56]
	Cauley 2010 (Hispanic)	0.57 [0.55, 0.59]
	Cauley 2010 (American indian)	0.61 [0.55, 0.67]
	Cauley 2010 (Asian)	0.58 [0.55, 0.61]
	Bolland 2011	0.60 [0.57, 0.63]
	Rubin 2013	0.70 [0.67, 0.74]
	Yu 2014 (Men)	0.61 [0.57, 0.65]
	Yu 2014 (Women)	0.60 [0.56, 0.63]
	van Geel 2014	0.65 [0.57, 0.74]
	Summary	**0.60 [0.57, 0.63]**
Garvan with BMD (10-year prediction) I^2=65.9%	Bolland 2011	0.64 [0.60, 0.67]
	Zhang 2011	0.70 [0.65, 0.75]
	Langsetmo 2011 (Men)	0.70 [0.65, 0.75]
	Langsetmo 2011 (Women)	0.69 [0.66, 0.72]
	Ahmed 2014 (Men)	0.61 [0.56, 0.67]
	Ahmed 2014 (Women)	0.62 [0.58, 0.65]
	van Geel 2014	0.69 [0.60, 0.78]
	Dominguez 2017	0.72 [0.61, 0.82]
	Summary	**0.67 [0.64, 0.69]**
Garvan without BMD (10-year prediction) I^2=69.9%	Sambrook 2011	0.64 [0.62, 0.66]
	Ahmed 2014 (Men)	0.57 [0.51, 0.63]
	Ahmed 2014 (Women)	0.58 [0.54, 0.61]
	van Geel 2014	0.65 [0.56, 0.73]
	Summary	**0.61 [0.57, 0.65]**
Garvan with BMD (5-year prediction) I^2=51.0%	Ahmed 2014 (Men)	0.67 [0.60, 0.74]
	Ahmed 2014 (Women)	0.61 [0.57, 0.65]
	Summary	**0.63 [0.57, 0.69]**
Garvan without BMD (5-year prediction) I^2=81.4%	Ahmed 2014 (Men)	0.67 [0.60, 0.74]
	Ahmed 2014 (Women)	0.57 [0.53, 0.61]
	Summary	**0.62 [0.52, 0.71]**
OSIRIS	Rubin 2013	0.68 [0.65, 0.72]
OST	Rubin 2013	0.68 [0.65, 0.72]
ORAI	Rubin 2013	0.69 [0.66, 0.72]
SCORE	Rubin 2013	0.69 [0.66, 0.73]
mSOF	Forti 2012	0.66 [0.57, 0.75]



prediction of hip fracture, but is considerably reduced in the absence of BMD. Calibration of the tools in different populations may also help in choosing the most appropriate tool. QFracture has been calibrated in the UK and Garvan in Australia. FRAX has been calibrated to the fracture and mortality rates of 62 countries in America, Europe, Africa, Asia, and Australia. Another unique feature of FRAX is that it was derived by taking account of the competing risk of death, an

Table 8 Summary AUC for prediction of Garvan-defined osteoporotic fracture, fragility fracture, or any fracture

Tools	Unadjusted model[e]		Adjusted model[e,f]	
	n	AUC [95% CI]	n	AUC [95% CI]
FRAX with BMD (10-year prediction)	10	0.64 [0.62, 0.66][b]	6	0.72 [0.65, 0.79][b,d]
FRAX without BMD (10-year prediction)	11	0.61 [0.58, 0.63][a,c]	6	0.69 [0.62, 0.76][a]
Garvan with BMD (10-year prediction)	8	0.66 [0.63, 0.69][b,d]	7	0.72 [0.66, 0.79][d]
Garvan without BMD (10-year prediction)	4	0.62 [0.58, 0.65][c]	3	0.67 [0.59, 0.74][a,c]

Garvan-defined osteoporotic fracture: hip, clinical spine, wrist, humerus, meta-carpal, scapula, clavicle, distal femur, tibia/fibula, patella, pelvis, or sternum

AUC area under the ROC curve, *CI* confidence interval, *n* number of AUCs

[a] Statistically different from AUC for FRAX with BMD

[b] Statistically different from AUC for FRAX without BMD

[c] Statistically different from AUC for Garvan with BMD

[d] Statistically different from AUC for Garvan without BMD

[e] Test for residual heterogeneity: $p < 0.0001$ (unadjusted model) and $p = 0.0033$ (adjusted model)

[f] Meta-regression model adjusted for gender, mean age, age range, and global quality. AUCs are presented for studies of high quality (score = 9/9) and cohort of women whose mean age is 65 years and age range is 40 years

approach widely recommended in the literature [83, 84]. With FRAX, fracture risk is thus reduced in individuals with low life expectancy, such as frailer and older people.

Our study has several limitations, but also numerous strengths. First, comparing performance of the different tools was an arduous task as discriminative accuracy is influenced by the characteristics and quality of the included studies. Meta-regression models were used to improve the comparability of the studies, but residual heterogeneity remained. Additional sources of heterogeneity could have been taken into account if the external validity of some tools such as Garvan and QFracture had been assessed in a larger number of studies. While calibration is an important attribute of prognostic models, calibration was not assessed. As gray literature was not searched, some studies may have been missed. When standard errors and confidence intervals for AUC were not reported, standard errors were approximated [85]. Since one of these measures was reported in most of the studies, this may only have affected the summary measures to a negligible extent. Despite these limitations, the results of our review are based on rigorous methodology. Studies published in all languages were searched. Selection, data abstraction, and quality assessment were performed by independent reviewers. All analyses were planned a priori. Our review is the first to use meta-regression models to compare the discriminative ability of current fracture risk prediction tools. This allowed us to perform formal statistical comparisons of the tools and to improve the comparability of included studies.

In conclusion, the results of this review suggest that QFracture is the tool with the highest discriminative ability to predict major osteoporotic fracture but that this tool is not significantly superior for prediction of hip fracture. Given the complexity of QFracture and the fact that it has only been validated in health electronic records, QFracture could be useful for automatic calculation of fracture risk and epidemiological research in electronic health records. For estimation of the risk hip fracture, FRAX with BMD and Garvan with BMD are among the best tools and are much easier to use in a clinical setting. However, when using these tools, information on BMD should be provided in order to avoid a significant reduction in discriminative ability. Given that the predictive ability is largely influenced by the characteristics of the validation studies, additional studies comparing the performance of current tools in the same cohort of individuals are needed.

Acknowledgements We gratefully aknowledge Vicky Tessier who has reviewed the search strategy.

Funding C Beaudoin has received a scholarship from the CHU de Québec and the Fonds de recherche du Québec-Santé (FRQS).

Compliance with ethical standards

Competing interests C Beaudoin, S Jean, L Moore, and M Gagné have no conflict of interest to disclose.

L Bessette has received grant/research support from Amgen Inc., BMS, Janssen, UCB, AbbVie, Pfizer, Sanofi, Eli Lilly, and Novartis; has consulted for Amgen Inc., BMS, Janssen, Roche, UCB, AbbVie, Pfizer, Merck, Celgene, Sanofi, Eli Lilly, and Novartis; and is a member of the Speakers' Bureau for Amgen Inc., BMS, Janssen, Roche, UCB, AbbVie, Pfizer, Merck, Celgene, Sanofi, Eli Lilly, and Novartis.

LG Ste-Marie has received grant/research support from Amgen Inc., has been a member of the advisory board of Amgen Inc. and Eli Lilly, and received other financial supports from AstraZeneca.

JP Brown has received grant/research support from Amgen Inc. and Eli Lilly; has consulted for Amgen Inc., Eli Lilly, and Merck; and is a member of the Speakers' Bureau for Amgen Inc. and Eli Lilly.

Appendix 1. Search strategy for Ovid MEDLINE

Table 9 Search Strategy for Ovid MEDLINE

Category	Searches	Results
Fracture	1. (Fracture* or (bone* adj2 (broken or break*))).ab,ti	1. 230,952
	2. exp *Fractures, Bone/	2. 138,448
	3. 1 or 2	3. 256,254
Risk prediction tool	4. 1 adj15 ((Risk* or Predict* or Prognostic* or Discrimin*) Adj4 (Tool* or Calculat* or Scor* or Instrument* or Model* or Index or Indexes or Indice* or Algorithm* or Nomogram* or Criteri* or Test* or Rule* or Scale*)).ab,ti	4. 2554
	5. exp Risk assessment/ and ((Tool* or Calculat* or Scor* or Instrument* or Model* or Index or Indexes or Indice* or Algorithm* or Nomogram* or Criteri* or Test* or Rule* or Scale*).ab,ti adj15 1)	5. 1532
	6. Risk Assessment/mt	6. 26,224
	7. (1 adj15 (Risk* or Predict* or Prognostic*).ab,ti) and (Models, Statistical/ or Models, biological/ or exp algorithms/ or exp Nomograms/ or exp "sensitivity and specificity"/ or exp roc curve/ or exp predictive value of tests/ or exp area under curve/ or validation studies.pt)	7. 2951
	8. exp Mass screening/	8. 121,866
	9. ((Decision or screen*) adj3 (Tool* or Calculat* or Scor* or Instrument* or Model* or Index or Indexes or Indice* or Algorithm* or Nomogram* or Criteri* or Test* or Rule* or Scale* or Approach* or Modalit* or Support* or Analys?s)).ab,ti	9. 142,975
	10. exp Decision support techniques/	10. 75,398
	11. 1 adj15 (Scor* adj (system* or method* or scheme)).ab,ti	11. 336
	12. (Intervention threshold*).ab,ti	12. 224
	13. 1 adj15 ((Risk* or probabilit*) adj4 (predict* or estim* or comput* or assess*)).ab,ti	13. 3957
	14. 1 adj15 (Risk* adj6 identif*).ab,ti	14. 1673
	15. (Fracture* Adj5 (Predict* or Discrimin*)).ab,ti	15. 4678
	16. 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15	16. 345,138
Known tools	17. (Abone or "Age body size no estrogen").ab,ti	17. 12
	18. (Bwc or weight criterion or pbw or patient body weight).ab,ti	18. 565
	19. (Caroc or ("Canadian Association of Radiologists" adj2 "Osteoporosis Canada")).ab,ti	19. 20
	20. Doescore.ab,ti	20. 2
	21. (Epese or ("established population*" adj3 "epidemiolog* stud*" adj3 elderly)).ab,ti	21. 486
	22. Epidos.ab,ti	22. 74
	23. (Fracture index or SOF or Study of osteoporotic fracture* OR Sofsurf).ab,ti	23. 2064
	24. (Framo or "fracture and mortality index").ab,ti	24. 4
	25. (Frax* or Fracture risk assessment).ab,ti	25. 3087
	26. (Frc OR Fracture risk calculator).ab,ti	26. 3341
	27. (Frisc or Frisk or (Fracture adj2 immobili?ation adj2 score) or Fracture risk score).ab,ti	27. 234
	28. Garvan*.ab,ti	28. 39
	29. (Gnudi S.au and Sitta E.au)	29. 6
	30. (KKOS* or Khon kaen osteoporosis study).ab,ti	30. 7
	31. (Male osteoporosis screening).ab,ti	31. 2
	32. (MORES OR Male osteoporosis risk estimation score).ab,ti	32. 414
	33. ((NOF or National osteoporosis foundation) adj3 guideline*).ab,ti	33. 66
	34. (OPERA or Osteoporotic prescreening risk assessment).ab,ti	34. 830
	35. (ORACLE or Osteoporosis Risk Assessment).ab,ti	35. 667
	36. (ORAI or Osteoporosis risk assessment instrument).ab,ti	36. 466
	37. (OSIRIS or Osteoporosis index).ab,ti	37. 269
	38. (OST or FOSTA or OSTA or Osteoporosis self-assessment or Osteoporosis screening tool).ab,ti	38. 1376
	39. Osteorisk.ab,ti	39. 4
	40. Qfracture*.ab,ti	40. 28
	41. (SCORE index or (Simple calculated adj2 risk estimation)).ab,ti	41. 1722
	42. 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41	42. 15,453
Studies	43. (Animals/ not Humans/)	43. 4,412,273
	44. Case reports.pt	44. 1,899,179
Combination	45. 3 and (16 or 42)	45. 13,650
	46. 45 not 43	46. 13,388
	47. 46 not 44	47. 13,189

Appendix 2. Quality assessment using a modified version of the QUADAS-2 tool

1. PATIENT SELECTION

    a. Risk of bias

- Was a consecutive (ex: administrative database) or random sample of individuals enrolled? *Yes, No, Unclear*
- Did the study avoid inappropriate exclusions? *Yes, No, Unclear*

    b. Concern regarding applicability

- Are there concerns that the included individuals and setting do not match the review question (Were participants from the validation cohort representative of individuals from the general population?)? *Low, High, Unclear*

2. INDEX TEST

    a. Concern regarding applicability

- Would the risk factors included in the index test would be available in clinical practice? *High, Low, Unclear*

3. REFERENCE STANDARD

    a. Risk of bias

- Is the reference standard likely to correctly classify the target condition (Were fractures verified or only self-reported?)? *Yes, No, Unclear*

4. FLOW AND TIMING

    a. Risk of bias

- Did all participants receive the same reference standard? *Yes, No, Unclear*
- Were at least 90% of eligible participants included in the analysis? *Yes, No, Unclear*
- Were more than 10% of missing data set to null or to the average value? *Yes, No, Unclear*
- Were all participants followed on the period for which the index test was constructed (Was the study duration (or maximum follow-up time) approximately the same as the length of the period for which the index test was constructed (±1 year)?)? *Yes, No, Unclear*
- Were more than 10 fractures by variable observed during the follow-up period? *Yes, No, Unclear*

Appendix 3. Summary of study quality as assessed using QUADAS-2 tool

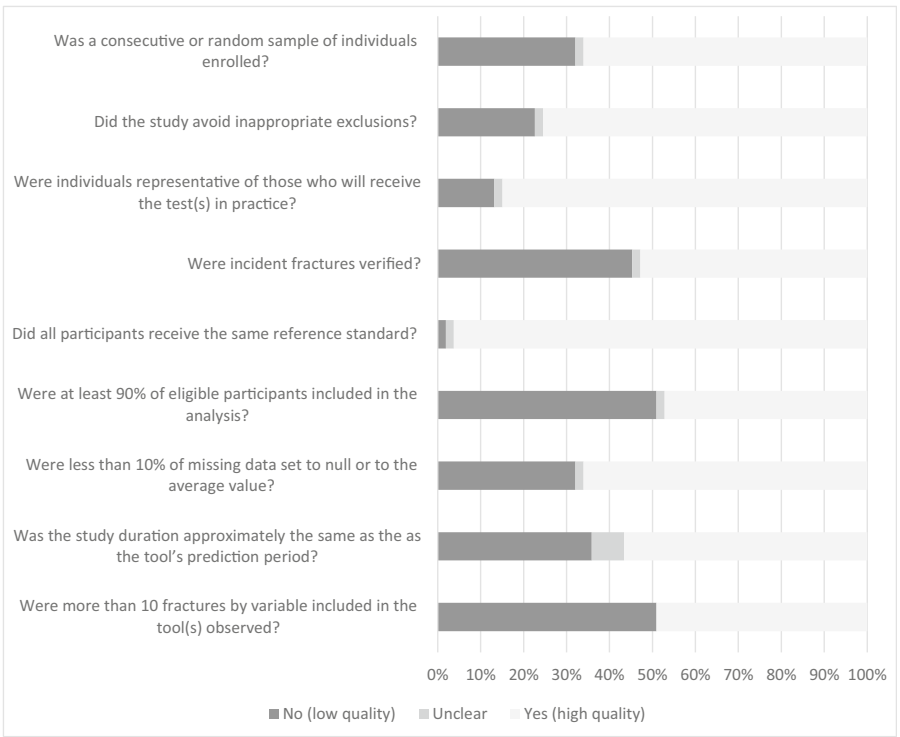


Fig. 2 Quality of the studies assessing the external validity of tools to identify individuals at high risk of non-traumatic fracture

References

1. NIH Consensus Development Panel on Osteoporosis Prevention, Diagnosis, and Therapy (2001) Osteoporosis prevention, diagnosis, and therapy. JAMA. 285(6):785–795
2. Osteoporosis Canada. Men and Osteoporosis [Available from: https://www.osteoporosis.ca/osteoporosis-and-you/men-and-osteoporosis/. Accessed 08 Mar 2019
3. Adachi JD, Loannidis G, Berger C, Joseph L, Papaioannou A, Pickard L, Papadimitropoulos EA, Hopman W, Poliquin S, Prior JC, Hanley DA, Olszynski WP, Anastassiades T, Brown JP, Murray T, Jackson SA, Tenenhouse A, Canadian Multicentre Osteoporosis Study (CaMos) Research Group (2001) The influence of osteoporotic fractures on health-related quality of life in community-dwelling men and women across Canada. Osteoporos Int 12(11): 903–908
4. Brenneman SK, Barrett-Connor E, Sajjan S, Markson LE, Siris ES (2006) Impact of recent fracture on health-related quality of life in postmenopausal women. J Bone Miner Res 21(6):809–816
5. Center JR, Nguyen TV, Schneider D, Sambrook PN, Eisman JA (1999) Mortality after all major types of osteoporotic fracture in men and women: an observational study. Lancet. 353(9156):878–882
6. Ioannidis G, Papaioannou A, Hopman WM, Akhtar-Danesh N, Anastassiades T, Pickard L, Kennedy CC, Prior JC, Olszynski WP, Davison KS, Goltzman D, Thabane L, Gafni A, Papadimitropoulos EA, Brown JP, Josse RG, Hanley DA, Adachi JD (2009) Relation between fractures and mortality: results from the Canadian multicentre osteoporosis study. CMAJ. 181(5):265–271
7. Haentjens P, Magaziner J, Colon-Emeric CS, Vanderschueren D, Milisen K, Velkeniers B et al (2010) Meta-analysis: excess mortality after hip fracture among older women and men. Ann Intern Med 152(6):380–390
8. Hopkins RB, Burke N, Von Keyserlingk C, Leslie WD, Morin SN, Adachi JD et al (2016) The current economic burden of illness of osteoporosis in Canada. Osteoporos Int 27(10):3023–3032
9. Cranney A, Jamal SA, Tsang JF, Josse RG, Leslie WD (2007) Low bone mineral density and fracture burden in postmenopausal women. CMAJ. 177(6):575–580
10. Compston J, Cooper A, Cooper C, Gittoes N, Gregson C, Harvey N et al (2017) UK clinical guideline for the prevention and treatment of osteoporosis. Arch Osteoporos 12(1):43
11. Rabar S, Lau R, O'Flynn N, Li L, Barry P (2012) Risk assessment of fragility fractures: summary of NICE guidance. BMJ. 345:e3698
12. Papaioannou A, Morin S, Cheung AM, Atkinson S, Brown JP, Feldman S, Hanley DA, Hodsman A, Jamal SA, Kaiser SM, Kvern B, Siminoski K, Leslie WD, for the Scientific Advisory Council of Osteoporosis Canada (2010) 2010 clinical practice guidelines for the diagnosis and management of osteoporosis in Canada: summary. CMAJ. 182(17):1864–1873
13. Marques A, Ferreira RJO, Santos E, Loza E, Carmona L, da Silva JAP (2015) The accuracy of osteoporotic fracture risk prediction tools: a systematic review and meta-analysis. Ann Rheum Dis 74(11):1958–1967
14. Rubin KH, Friis-Holmberg T, Hermann AP, Abrahamsen B, Brixen K (2013) Risk assessment tools to identify women with increased risk of osteoporotic fracture: complexity or simplicity? A systematic review. J Bone Miner Res 28(8):1701–1717
15. Nayak S, Edwards DL, Saleh AA, Greenspan SL (2014) Performance of risk assessment instruments for predicting osteoporotic fracture risk: a systematic review. Osteoporos Int 25(1):23–49
16. Leslie WD, Lix LM (2014) Comparison between various fracture risk assessment tools. Osteoporos Int 25(1):1–21
17. Iki M, Fujita Y, Tamaki J, Kouda K, Yura A, Sato Y, Moon JS, Winzenrieth R, Okamoto N, Kurumatani N (2015) Trabecular bone score may improve FRAX prediction accuracy for major osteoporotic fractures in elderly Japanese men: the Fujiwara-kyo osteoporosis risk in men (FORMEN) cohort study. Osteoporos Int 26(6): 1841–1848
18. Lundin H, Torabi F, Saaf M, Strender LE, Nyren S, Johansson SE et al (2015) Laser-supported dual energy X-ray absorptiometry (DXL) compared to conventional absorptiometry (DXA) and to FRAX as tools for fracture risk assessments. PLoS One 10(9): e0137535
19. Azagra R, Zwart M, Encabo G, Aguye A, Martin-Sanchez JC, Puchol-Ruiz N et al (2016) Rationale of the Spanish FRAX model in decision-making for predicting osteoporotic fractures: an update of FRIDEX cohort of Spanish women. BMC Musculoskelet Disord 17(1):262
20. Goldshtein I, Gerber Y, Ish-Shalom S, Leshno M (2016) Validation of fracture risk assesment tool using real-world data. Pharmacoepidemiol Drug Saf 25:250–251
21. Holloway KL, Mohebbi M, Hans D, Brennan-Olsen SL, Kotowicz MA, Pasco JA (2016) Prediction of hip fractures in Australian men using FRAX scores adjusted with trabecular bone score. Osteoporos Int 27(1 SUPPL. 1):S351
22. Klop C, De Vries F, Bijlsma JWJ, Leufkens HGM, Welsing PMJ (2016) Predicting the 10-year risk of hip and major osteoporotic fracture in rheumatoid arthritis and in the general population: an independent validation and update of UK FRAX without bone mineral density. Ann Rheum Dis 75(12):2095–2100
23. Sundh V, Jonasson G (2016) FRAX and mandibular sparse trabeculation as fracture predictors: a longitudinal study 1980–2002. Osteoporos Int 27(1 SUPPL. 1):S257
24. Dagan N, Cohen-Stavi C, Leventer-Roberts M, Balicer RD (2017) External validation and comparison of three prediction tools for risk of osteoporotic fractures using data from population based electronic health records: retrospective cohort study. BMJ (Online) 356: i6755
25. Francesco L, Elisa B, Raffaella M, Alessandro P, Iacopo C, Giampiero M, et al. (2017) Assessing risk of osteoporotic fractures in primary care: development and validation of the FRA-HS algorithm. Calcif Tissue Int 1–13
26. Hoff M, Meyer HE, Skurtveit S, Langhammer A, Sogaard AJ, Syversen U et al (2017) Validation of FRAX and the impact of self-reported falls among elderly in a general population: the HUNT study. Norway Osteoporos Int:1–10
27. Olmos JM, Hernandez JL, Gonzalez JL, Martinez J, Pariente E, Sierra I et al (2017) Predictive and discriminatory capacity of the frax tool in Spanish postmenopathic women: a preliminary study. Calcif Tissue Int 100(1 Supplement 1):S86
28. Orwoll ES, Lapidus J, Wang PY, Vandenput L, Hoffman A, Fink HA, Laughlin GA, Nethander M, Ljunggren Ö, Kindmark A, Lorentzon M, Karlsson MK, Mellström D, Kwok A, Khosla S, Kwok T, Ohlsson C, for the Osteoporotic Fractures in Men (MrOS) Study Research Group (2017) The limited clinical utility of testosterone, estradiol, and sex hormone binding globulin measurements in the prediction of fracture risk and bone loss in older men. J Bone Miner Res 32(3):633–640
29. Reyes Dominguez AI, Sosa Cabrera N, Saavedra Santana P, de Tejada Romero MJG, Jodar Gimeno E, Sosa HM (2017) Assessment of the predictive capacity of the garvan calculator of 10 year risk of fracture in a Spanish population. Revista de Osteoporosis y Metabolismo Mineral 9(2):55–61
30. Su Y, Leung J, Hans D, Lamy O, Kwok T (2017) The added value of trabecular bone score to FRAX to predict major osteoporotic fractures for clinical use in Chinese older people: the Mr. OS and Ms. OS cohort study in Hong Kong. Osteoporos Int 28(1):111–117

31. Deeks JJ, Bossuyt PM, Gatsonis C (editors) (2013). Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 1.0.0 [updated September 2013]: The Cochrane Collaboration. Available from: http://srdta.cochrane.org/. Accessed 17 Mar 2014
32. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ. 339:b2535
33. Toll DB, Janssen KJ, Vergouwe Y, Moons KG (2008) Validation, updating and impact of clinical prediction rules: a review. J Clin Epidemiol 61(11):1085–1094
34. Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 143(1):29–36
35. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM, QUADAS-2 Group (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155(8):529–536
36. Bennett DA (2001) How can I deal with missing data in my study? Aust N Z J Public Health 25(5):464–469
37. Peduzzi P, Concato J, Feinstein AR, Holford TR (1995) Importance of events per independent variable in proportional hazards regression analysis. II. Accuracy and precision of regression estimates. J Clin Epidemiol 48(12):1503–1510
38. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 49(12):1373–1379
39. Viechtbauer W (2010) Conducting meta-analyses in R with the metafor Package. J Stat Softw 36(3):48
40. Hosmer DW, Lemeshow S (1989) Applied logistic regression, vol xiii. Wiley, New York, p 307
41. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]: The Cochrane Collaboration; 2011. Available from: http://handbook.cochrane.org/. Accessed 02 Mar 2014
42. Pressman AR, Lo JC, Chandra M, Ettinger B (2011) Methods for assessing fracture risk prediction models: experience with FRAX in a large integrated health care delivery system. J Clin Densitom 14(4):407–415
43. Lo JC, Pressman AR, Chandra M, Ettinger B (2011) Fracture risk tool validation in an integrated healthcare delivery system. Am J Manag Care 17(3):188–194
44. Ettinger B, Liu H, Blackwell T, Hoffman AR, Ensrud KE, Orwoll ES, Osteoporotic Fracture in Men (MrOS) Research Group (2012) Validation of FRC, a fracture risk assessment tool, in a cohort of older men: the osteoporotic fractures in men (MrOS) study. J Clin Densitom 15(3):334–342
45. Ettinger B, Ensrud KE, Blackwell T, Curtis JR, Lapidus JA, Orwoll ES et al (2013) Performance of FRAX in a cohort of community-dwelling, ambulatory older men: the osteoporotic fractures in men (MrOS) study. Osteoporos Int 24(4):1185–1193
46. Yu R, Leung J, Woo J (2014) Sarcopenia combined with FRAX probabilities improves fracture risk prediction in older Chinese men. J Am Med Dir Assoc 15(12):918–923
47. Tebe Cordomi C, Del Rio LM, Di Gregorio S, Casas L, Estrada M-D, Kotzeva A et al (2013) Validation of the FRAX predictive model for major osteoporotic fracture in a historical cohort of Spanish women. J Clin Densitom 16(2):231–237
48. Czerwinski E, Borowy P, Kumorek A, Amarowicz J, Gorkiewicz M, Milert A (2013) Fracture risk prediction in outpatients from Krakow region using FRAX tool versus fracture risk in 11-year follow-up. Ortopedia, traumatologia, rehabilitacja 15(6):617–628
49. Sornay-Rendu E, Munoz F, Delmas PD, Chapurlat RD (2010) The FRAX tool in French women: how well does it describe the real incidence of fracture in the OFELY cohort? J Bone Miner Res 25(10):2101–2107
50. Cheung EYN, Bow CH, Cheung CL, Soong C, Yeung S, Loong C, Kung A (2012) Discriminative value of FRAX for fracture prediction in a cohort of Chinese postmenopausal women. Osteoporos Int 23(3):871–878
51. Tremollieres FA, Pouilles J-M, Drewniak N, Laparra J, Ribot CA, Dargent-Molina P (2010) Fracture risk prediction using BMD and clinical risk factors in early postmenopausal women: sensitivity of the WHO FRAX tool. J Bone Miner Res 25(5):1002–1009
52. Yun H, Delzell E, Ensrud KE, Kilgore ML, Becker D, Morrisey MA, Curtis JR (2010) Predicting hip and major osteoporotic fractures using administrative data. Arch Intern Med 170(21):1940–1942
53. Gonzalez-Macias J, Marin F, Vila J, Diez-Perez A (2012) Probability of fractures predicted by FRAX and observed incidence in the Spanish ECOSAP study cohort. Bone. 50(1):373–377
54. Tamaki J, Iki M, Kadowaki E, Sato Y, Kajita E, Kagamimori S, Kagawa Y, Yoneshima H (2011) Fracture risk prediction using FRAX: a 10-year follow-up survey of the Japanese population-based osteoporosis (JPOS) cohort study. Osteoporos Int 22(12):3037–3045
55. Briot K, Paternotte S, Kolta S, Eastell R, Felsenberg D, Reid DM, Glüer CC, Roux C (2013) FRAX: prediction of major osteoporotic fractures in women from the general population: the OPUS study. PLoS One 8(12):e83436
56. Ensrud KE, Lui L-Y, Taylor BC, Schousboe JT, Donaldson MG, Fink HA, Cauley JA, Hillier TA, Browner WS, Cummings SR, Study of Osteoporotic Fractures Research Group (2009) A comparison of prediction models for fractures in older women: is more better? Arch Intern Med 169(22):2087–2094
57. Rubin KH, Abrahamsen B, Friis-Holmberg T, Hjelmborg JVB, Bech M, Hermann AP, Barkmann R, Glüer CC, Brixen K (2013) Comparison of different screening tools (FRAX, OST, ORAI, OSIRIS, SCORE and age alone) to identify women with increased risk of fracture. A population-based prospective study. Bone. 56(1):16–22
58. Langsetmo L, Nguyen TV, Nguyen ND, Kovacs CS, Prior JC, Center JR, Morin S, Josse RG, Adachi JD, Hanley DA, Eisman JA, the Canadian Multicentre Osteoporosis Study Research Group (2011) Independent external validation of nomograms for predicting risk of low-trauma fracture and hip fracture. CMAJ. 183(2):E107–E114
59. Bolland MJ, Siu AT, Mason BH, Horne AM, Ames RW, Grey AB et al (2011) Evaluation of the FRAX and Garvan fracture risk calculators in older women. J Bone Miner Res 26(2):420–427
60. Henry MJ, Pasco JA, Merriman EN, Zhang Y, Sanders KM, Kotowicz MA, Nicholson GC (2011) Fracture risk score and absolute risk of fracture. Radiology. 259(2):495–501
61. Zhang Y, Pasco JA, Kotowicz MA, Sanders KM, Nicholson GC, Henry MJ (2011) How well do the frax (AUS) and garvan calculators predict fractures from the Geelong Osteoporosis Study (GOS). IOF Regionals, 2nd Asia-Pacific Osteoporosis and Bone Meeting, ANZBMS Annual Scientific Meeting, with JSBMR Gold Coast, QLD Australia: Springer London. S548
62. Leslie WD, Lix LM, Johansson H, Oden A, McCloskey E, Kanis JA, Manitoba Bone Density Program (2010) Independent clinical validation of a Canadian FRAX tool: fracture prediction and model calibration. J Bone Miner Res 25(11):2350–2358
63. Morin S, Tsang JF, Leslie WD (2009) Weight and body mass index predict bone mineral density and fractures in women aged 40 to 59 years. Osteoporos Int 20(3):363–370
64. Hundrup YA, Jacobsen RK, Andreasen AH, Davidsen M, Obel EB, Abrahamsen B (2010) Validation of a 5-year risk score of hip fracture in postmenopausal women. The Danish nurse cohort study. Osteoporos Int 21(12):2135–2142
65. Hundrup YA, Jacobsen RK, Andreasen AH, Davidsen M, Obel EB, Abrahamsen B. Performance of BMD-independent risk scores

adapted from the WHI and FRAX(R) collaborations in prediction of 5-year risk of hip fracture: The Danish nurses cohort. 36th European Symposium on Calcified Tissues; Vienna Austria: Elsevier Inc.; 2009. p. S397
66. Albaba M, editor (2014) Is HIP fracture risk assessment index (HFRAI), an electronic medical database derived tool, comparable to the world health organization fracture assessment tool (FRAX) in subjects without known femoral neck bone mineral density? 2014 Annual scientific meeting of the American Geriatrics Society. Orlando: Blackwell Publishing Inc.
67. Cauley J, LaCroix AZ, Wu C, Lewis B, Wactawski-Wende J, Masaki K et al (2010) FRAX: does fracture prediction differ by race/ethnicity? 32nd annual meeting of the American Society for Bone and Mineral Research. Wiley-Blackwell, Toronto, ON Canada
68. Robbins J, Aragaki AK, Kooperberg C, Watts N, Wactawski-Wende J, Jackson RD, LeBoff MS, Lewis CE, Chen Z, Stefanick ML, Cauley J (2007) Factors associated with 5-year risk of hip fracture in postmenopausal women. JAMA. 298(20):2389–2398
69. Ahmed LA, Nguyen ND, Bjornerem A, Joakimsen RM, Jorgensen L, Stormer J et al (2014) External validation of the garvan nomograms for predicting absolute fracture risk: the tromso study. PLoS One 9:e107695
70. Tanaka S, Yoshimura N, Kuroda T, Hosoi T, Saito M, Shiraki M (2010) The fracture and immobilization score (FRISC) for risk assessment of osteoporotic fracture and immobilization in postmenopausal women—a joint analysis of the Nagano, Miyama, and Taiji cohorts. Bone. 47(6):1064–1070
71. Hippisley-Cox J, Coupland C, Brindle P (2014) The performance of seven QPrediction risk scores in an independent external sample of patients from general practice: a validation study. BMJ Open 4(8):e005809
72. Collins GS, Mallett S, Altman DG (2011) Predicting risk of osteoporotic and hip fracture in the United Kingdom: prospective independent and external validation of QFractureScores. BMJ (Clinical research ed) 342:d3651
73. Hippisley-Cox J, Coupland C (2009) Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores. BMJ (Clinical research ed) 339:b4229
74. van Geel TACM, Eisman JA, Geusens PP, van den Bergh JPW, Center JR, Dinant G-J (2014) The utility of absolute risk prediction using FRAX and Garvan fracture risk calculator in daily practice. Maturitas. 77(2):174–179
75. Friis-Holmberg T, Rubin KH, Brixen K, Tolstrup JS, Bech M (2014) Fracture risk prediction using phalangeal bone mineral density or FRAX()?-a Danish cohort study on men and women. J Clin Densitom 17(1):7–15
76. Forti P, Rietti E, Pisacane N, Olivelli V, Maltoni B, Ravaglia G (2012) A comparison of frailty indexes for prediction of adverse health outcomes in an elderly cohort. Arch Gerontol Geriatr 54(1): 16–20
77. Tanaka S, Kuroda T, Saito M, Shiraki M (2011) Urinary pentosidine improves risk classification using fracture risk assessment tools for postmenopausal women. J Bone Miner Res 26(11):2778–2784
78. Sambrook PN, Flahive J, Hooven FH, Boonen S, Chapurlat R, Lindsay R, Nguyen TV, Díez-Perez A, Pfeilschifter J, Greenspan SL, Hosmer D, Netelenbos JC, Adachi JD, Watts NB, Cooper C, Roux C, Rossini M, Siris ES, Silverman S, Saag KG, Compston JE, LaCroix A, Gehlbach S (2011) Predicting fractures in an international cohort using risk factor algorithms without BMD. J Bone Miner Res 26(11):2770–2777
79. Sund R, Honkanen R, Johansson H, Oden A, McCloskey E, Kanis J et al (2014) Evaluation of the FRAX model for hip fracture predictions in the population-based Kuopio osteoporosis risk factor and prevention study (OSTPRE). Calcif Tissue Int 95(1):39–45
80. Albertsson DM, Mellstrom D, Petersson C, Eggertsen R (2007) Validation of a 4-item score predicting hip fracture and mortality risk among elderly women. Ann Fam Med 5(1):48–56
81. Beaudoin C, Jean S, Moore L, Gamache P, Bessette L, Ste-Marie LG et al (2018) Number, location, and time since prior fracture as predictors of future fracture in the elderly from the general population. J Bone Miner Res 33(11):1956–1966
82. Johansson H, Kanis JA, Oden A, McCloskey E, Chapurlat RD, Christiansen C et al (2014) A meta-analysis of the association of fracture risk and body mass index in women. J Bone Miner Res 29(1):223–233
83. Lau B, Cole SR, Gange SJ (2009) Competing risk regression models for epidemiologic data. Am J Epidemiol 170(2):244–256
84. Leslie WD, Lix LM, Wu X (2013) Competing mortality and fracture risk assessment. Osteoporos Int 24(2):681–688
85. Hanley JA, McNeil BJ (1983) A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology. 148(3):839–843