Osteoporosis International (2019) 30:2129–2139
https://doi.org/10.1007/s00198-019-05088-2

ORIGINAL ARTICLE

# How well do the FRAX (Australia) and Garvan calculators predict incident fractures? Data from the Geelong Osteoporosis Study

K. L. Holloway-Kew[1] · Y. Zhang[2] · A. G. Betson[1] · K. B. Anderson[1] · D. Hans[3] · N. K. Hyde[1] · G. C. Nicholson[4] · N. A. Pocock[5] · M. A. Kotowicz[1,3,6] · J. A. Pasco[1,3,6]

Received: 18 November 2018 / Accepted: 9 July 2019 / Published online: 18 July 2019


## Abstract

**Summary** This study reports that both FRAX and Garvan calculators underestimated fractures in Australian men and women, particularly in those with osteopenia or osteoporosis. Major osteoporotic fractures were poorly predicted, while both calculators performed acceptably well for hip fractures.

**Introduction** This study assessed the ability of the FRAX (Australia) and Garvan calculators to predict fractures in Australian women and men.

**Methods** Women ($n = 809$) and men ($n = 821$) aged 50–90 years, enrolled in the Geelong Osteoporosis Study, were included. Fracture risk was estimated using FRAX and Garvan calculators with and without femoral neck bone mineral density (BMD) ($FRAX_{BMD}$, $FRAX_{noBMD}$, $Garvan_{BMD}$, $Garvan_{noBMD}$). Incident major osteoporotic (MOF), fragility, and hip fractures over the following 10 years were verified radiologically. Differences between observed and predicted numbers of fractures were assessed using a chi-squared test. Diagnostics indexes were calculated.

**Results** In women, 115 MOF, 184 fragility, and 42 hip fractures occurred. For men, there were 73, 109, and 17 fractures, respectively. FRAX underestimated MOFs, regardless of sex or inclusion of BMD. FRAX accurately predicted hip fractures, except in women with BMD (20 predicted, $p = 0.004$). Garvan underestimated fragility fractures except in men using BMD (88 predicted, $p = 0.109$). Garvan accurately predicted hip fractures except for women without BMD (12 predicted, $p < 0.001$). Fractures were underestimated primarily in the osteopenia and osteoporosis groups; MOFs in the normal BMD group were only underestimated by $FRAX_{BMD}$ and fragility fractures by $Garvan_{noBMD}$, both in men. AUROCs were not different between scores with and without BMD, except for fragility fractures predicted by Garvan in women (0.696, 95% CI 0.652–0.739 and 0.668, 0.623–0.712, respectively, $p = 0.008$) and men, which almost reached significance (0.683, 0.631–0.734, and 0.667, 0.615–0.719, respectively, $p = 0.051$). Analyses of sensitivity and specificity showed overall that MOFs and fragility fractures were poorly predicted by both FRAX and Garvan, while hip fractures were acceptably predicted.

**Conclusions** Overall, the FRAX and Garvan calculators underestimated MOF and fragility fractures, particularly in individuals with osteopenia or osteoporosis. Hip fractures were predicted better by both calculators. AUROC analyses suggest that $Garvan_{BMD}$ performed better than $Garvan_{noBMD}$ for prediction of fragility fractures.

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s00198-019-05088-2) contains supplementary material, which is available to authorized users.

✉ K. L. Holloway-Kew
khollo@barwonhealth.org.au

[1] School of Medicine, Deakin University, Geelong, Australia

[2] Department of Medicine-Western Health, Melbourne Medical School, The University of Melbourne, St Albans, Australia

[3] Center of Bone Diseases, Bone & Joint Department, Lausanne University Hospital, Lausanne, Switzerland

[4] Rural Clinical School, The University of Queensland, Toowoomba, QLD, Australia

[5] University of New South Wales, Sydney, NSW, Australia

[6] Barwon Health, Geelong, Australia

**Keywords** Absolute fracture risk calculator · FRAX · Garvan · Geelong osteoporosis study

## Introduction

Fragility fractures are associated with increased morbidity and mortality, particularly among postmenopausal women and elderly men [1–4]. There are several anti-fracture interventions available, and these should directed to those at higher risk, such as women with osteopenia aged over 65 years, to improve efficacy and cost-effectiveness [5–7]. Prognostic models facilitate clinical decision-making by attempting to distinguish those with higher risk of fracture from those at lower risk.

FRAX [8, 9] is a web-based calculator designed to estimate fracture risk, developed by the University of Sheffield, UK, using fracture and risk data from multi-national cohorts to derive country-specific calculators for absolute fracture risk. It was designed to predict the 10-year absolute risk of 'major osteoporotic' fractures (MOF; hip, clinical spine, forearm, and proximal humerus) by combining up to 11 risk factors including femoral neck bone mineral density (BMD), age, sex, weight, height, previous fracture, parental hip fractures, smoking, glucocorticoid usage, rheumatoid arthritis, secondary osteoporosis, and alcohol consumption. FRAX was first available in 2008 [10] and has been designed to be applied with or without BMD measurements, as access to densitometers is sometimes limited. The FRAX tool takes into account mortality as a competing event with fracture when calculating the scores. FRAX is intended to guide clinical treatment decisions for patients aged ≥ 50 years with a BMD T-score between − 1.0 and − 2.5. If a patient has a FRAX 10-year probability of ≥ 3% for hip or ≥ 20% for MOF, they are considered to be at high risk and anti-fracture treatment is recommended [11].

The Garvan calculator was derived using fracture and risk factor data from the Dubbo Osteoporosis Epidemiology Study (DOES) [12, 13]. It predicts 5- and 10-year absolute risk of hip and "fragility" fractures (low trauma fractures excluding cervical spine, digits, and skull/face) using the risk factors of age, sex, past falls, previous low trauma fractures occurring after the age of 50 years and femoral neck BMD. The Garvan calculator uses weight (kg) when femoral neck BMD measurement is unavailable. The Garvan calculator does not take into account competing mortality when calculating fracture risk scores.

Previous studies have assessed the discriminatory abilities of the FRAX and Garvan calculators by observing area under receiver operating characteristic (ROC) curves [14–17] and reported no significant difference. Kanis et al. [18] have pointed out the limitations of using ROCs, including (1) poor sensitivity when analysing additional variables that are of low prevalence. For example, addition of important fracture risk factors such as smoking would have little impact on the area under the AUROC curves if the prevalence of smoking among the study sample is low; (2) comparisons of ROC across different studies lack consistency as the sample demographics of each study may be different; and (3) the ROC curves are inappropriate for the determination of an intervention threshold. New methods of comparison, therefore, should be explored for the external validation of the fracture risk calculators. The best method is comparison of the number of fractures observed with the number predicted by the fracture risk calculator [18]. Several studies have investigated the number of fractures predicted by both FRAX and Garvan calculators and the actual number of fractures observed, with some reporting good prediction of fractures, while others reported under- or over-estimation of fractures [15, 17, 19, 20].

A fracture assessment calculator should primarily target the population from which it was derived, as fracture and mortality hazards may vary between populations [18]. The Australian version of FRAX and Garvan were derived using data from the Australian Dubbo Osteoporosis Epidemiology Study. Our aim was to validate the FRAX (Australia) and Garvan calculators in a different Australian cohort.

## Methods

### Study region and sample cohort

This study was set in the Barwon Statistical Division (BSD), as part of the Geelong Osteoporosis Study (GOS) [21]. The BSD is situated in south-eastern Australia and has a population of ~ 280,000. The region resembles the broader Australian population and has centralised health services in the regional city of Geelong, making it an ideal site for epidemiological research.

An age-stratified random sample of 1494 women aged 20–94 years was recruited from the Commonwealth electoral roll during 1993–1997 [21]. A cohort of men ($n$ = 1540) aged 20–96 years was similarly recruited during 2001–2006. Participants aged 50–90 years were included in this study, as this covers the age range routinely assessed for fracture risk in a clinical setting.

### Fracture ascertainment

Participants were followed prospectively either for 10 years, until date of first fracture, or date of death. Participants were followed by examination of radiological records. Incident fractures that occurred during follow-up were identified

and confirmed radiologically using x-ray reports from all radiological centres servicing the region. This method of fracture ascertainment has been previously validated [22]. Trained research staff individually examined each record and determined the most appropriate ICD-9 (international code of diseases version 9) codes for fracture site, as well as level of trauma (further details below). Codes were not directly extracted from medical records. Clinical vertebral fractures were classified as those which presented clinically, and had a 25% or more reduction in vertebral height [23]. Serial X-rays were not performed as part of this study, and thus non-clinical vertebral fractures may not have been detected. Only the first incident hip or MOF (for FRAX) and first incident hip or fragility fracture (for Garvan) during follow-up was considered for each analysis. Information regarding the cause, anatomical sites and date of the incident fracture was recorded using ICD-9 codes. The following fracture sites were considered for inclusion in FRAX fracture risk calculations [9]:

- ICD9 code 820 (hip)
- ICD9 code 805 (clinical spine)
- ICD9 code 813.4 (distal forearm)
- ICD9 code 812 (proximal humerus only)

For Garvan, all *except* the following fracture sites were included [12]:

- ICD9 code 802 (face)
- ICD9 code 803 (skull)
- ICD9 code 805 (only cervical spine was excluded)
- ICD9 code 816 (finger)
- ICD9 code 826 (toe)

The cause of fracture was also classified according to ICD-9 codes. Only fractures with the following codes corresponding to minimal trauma were included (for both FRAX and Garvan):

- ICD9 code 885 (fall from slipping, tripping and or stumbling on the same level)
- ICD9 code 886 (fall on the same level from collision, pushing, shoving, by or with another person)
- ICD9 code 887 (unspecific cause)
- ICD9 code 888 (other and unspecific fall)
- ICD9 code 927 (overexertion and strenuous movements)

There were 46 fractures in women and 15 fractures in men that were excluded due to high trauma. Pathological fractures were also excluded; however, very few of these occurred during the study ($n = 5$). Deaths were identified by data linkage with the National Deaths Index (Australian Institute of Health and Welfare).

The study was approved by the Barwon Health Human Research Ethics Committee, and all participants gave written informed consent.

## Measurements

At the baseline visit for women, BMD at the femoral neck was measured using a Lunar DPX-L densitometer (software version 1.31); short-term precision in vivo at this site was 1.6%. For men at baseline, the first 544 men were scanned on the Lunar DPX-L, and when the DPX-L became outmoded, scans were performed using a GE-Prodigy (Prodigy; GE Lunar, Madison, WI, USA). No differences were detected in BMD at the lumbar spine or femoral neck when cross-calibration was performed [21]. Participants' weight and height were measured to the nearest 0.1 kg and 0.1 cm. Prior fractures, parental history of hip fractures, rheumatoid arthritis, smoking, use of glucocorticoids, and secondary osteoporosis were determined by a questionnaire administered by trained interviewers. Secondary osteoporosis included type 1 diabetes, osteogenesis imperfecta, hyperthyroidism, hypogonadism, menopause before the age of 45 years, malabsorption, or chronic liver disease. Alcohol consumption was determined using a validated Food Frequency Questionnaire [24]. Number of falls were self-reported over the previous 12 months.

Ten-year fracture risk estimates for both FRAX (Australia) and Garvan were obtained by entering the appropriate data into the online calculators. These are referred to in the text as FRAX or Garvan MOF, fragility, and hip fracture scores as appropriate. BMD values were directly entered for both FRAX and Garvan, and T-scores were automatically calculated by the calculators. Four FRAX fracture risk scores were calculated per participant: (1) major osteoporotic fracture FRAX with BMD (MOF $FRAX_{BMD}$), (2) major osteoporotic fracture FRAX without BMD (MOF $FRAX_{noBMD}$), (3) hip fracture FRAX with BMD (Hip $FRAX_{BMD}$), and (4) hip Fracture FRAX without BMD (Hip $FRAX_{noBMD}$).

Four Garvan scores were calculated per participant: (1) fragility fracture Garvan with BMD (Fragility $Garvan_{BMD}$), (2) fragility fracture Garvan without BMD (Fragility $Garvan_{noBMD}$), (3) hip fracture Garvan with BMD (Hip $Garvan_{BMD}$), and (4) hip fracture Garvan without BMD (Hip $Garvan_{noBMD}$).

## Statistical analysis

The two-sample *t* test or Mann-Whitney test was used to determine differences for continuous variables according to fracture status. A chi-square test or Fisher's test was used for categorical variables. For both fracture risk calculators, absolute risk was calculated with and without femoral neck BMD measurements and expressed as a percentage. The predicted

number of fractures for each score was determined by multiplying the absolute risk by the proportion of follow-up. The difference between the total predicted and observed number of fractures was assessed with a chi-squared test. No significant difference between the predicted and observed number would indicate that the calculators were effective at estimating fracture risk. This analysis was completed in two ways: (1) the entire participant group and (2) with participants divided into categories based on femoral neck BMD T-scores that were calculated using Australian reference ranges [25, 26]. This was completed to determine how well the calculators predicted fractures in low and high-risk groups.

The areas under receiver operating characteristics (AUROC) curves were quantified for $FRAX_{noBMD}$ and $FRAX_{BMD}$, as well as $Garvan_{noBMD}$ and $Garvan_{BMD}$. The AUROC provides a measure of ability to discriminate independent of a particular threshold, with higher values corresponding to improved predictions: < 0.70, poor; 0.70–0.79, fair; 0.80–0.89, good; and 0.90–1.00, excellent [27]. AUROCs were then compared using a test for the equality of the area under the curves, using an algorithm suggested by DeLong et al. [28]. Diagnostic indexes including sensitivity, specificity, correctly classified, likelihood ratios, and predictive values were calculated for all fracture risk scores. For FRAX, 10-year probability cut-points of ≥ 20% and ≥ 3% were used for MOF and hip fracture respectively, according to US National Osteoporosis Foundation guidelines [11], as there are no specific recommendations in the Australian context. For Garvan, 10-year probability cut-points of ≥ 14% and ≥ 3% were used for fragility and hip fracture [29]. As a general rule, both sensitivity and specificity of ≥ 80% is considered a good model for predictions. An acceptable model is considered to have both specificity and sensitivity of 50–79%. Sensitivity or specificity lower than 50% is considered a poor model.

Analyses were performed using the statistical packages Minitab version 18 and STATA version 15.

## Results

### Descriptive characteristics

There were 836 women aged 50–90 years at baseline. Of these, 13 did not have a BMD measurement, four did not have sufficient information to calculate FRAX and five did not have

**Table 1** Participant descriptive characteristics, stratified by major osteoporotic fracture (hip, clinical spine, forearm, and proximal humerus) status. Data are expressed as mean (SD), median (IQR), or *n* (%)

| | Women | | | | Men | | | |
|---|---|---|---|---|---|---|---|---|
| | All (*n* = 809) | Fracture (*n* = 115) | No fracture (*n* = 694) | *p* | All (*n* = 821) | Fracture (*n* = 73) | No fracture (*n* = 748) | *p* |
| Age (years) | 71.0 (60.0–72.0) | 72.0 (71.0–81.0) | 69.5 (59.0–72.0) | *< 0.001* | 69.0 (59.0–78.0) | 77.0 (70.5–83.0) | 68.0 (58.0–77.0) | *< 0.001* |
| Weight (kg) | 67.6 ± 13.3 | 63.9 ± 11.9 | 68.2 ± 13.4 | *0.001* | 82.0 ± 13.0 | 80.8 ± 11.0 | 82.2 ± 13.2 | 0.314 |
| Height (cm) | 158.1 ± 6.4 | 156.3 ± 6.2 | 158.3 ± 6.4 | *0.002* | 172.8 ± 6.6 | 171.7 ± 6.1 | 172.9 ± 6.7 | 0.118 |
| Prior fracture (yes) | 168 (20.8) | 41 (35.7) | 127 (18.3) | *< 0.001* | 157 (19.1) | 22 (30.1) | 135 (18.1) | *0.012* |
| Parent fractured hip (yes) | 59 (7.3) | 9 (7.8) | 50 (7.2) | 0.812 | 52 (6.3) | 3 (4.1) | 49 (6.6) | 0.414 |
| Current smoking (yes) | 75 (9.3) | 9 (7.8) | 66 (9.5) | 0.564 | 80 (9.7) | 7 (9.6) | 73 (9.8) | 0.963 |
| Glucocorticoid use (yes) | 25 (3.1) | 10 (8.7) | 15 (2.2) | *< 0.001* | 20 (2.4) | 7 (9.6) | 13 (1.7) | *< 0.001* |
| Rheumatoid arthritis (yes) | 119 (14.7) | 23 (20.0) | 96 (13.8) | 0.084 | 51 (6.2) | 8 (11.0) | 43 (5.8) | 0.078 |
| Secondary osteoporosis (yes) | 159 (19.7) | 30 (26.1) | 129 (18.6) | 0.061 | 16 (2.0) | 2 (2.7) | 14 (1.9) | 0.609 |
| Alcohol ≥ 3 units/day (yes) | 6 (0.7) | 0 (0.0) | 6 (0.9) | † | 189 (23.0) | 15 (20.6) | 174 (23.3) | 0.599 |
| Falls (last 12 months) (yes) | 165 (20.4) | 32 (27.8) | 133 (19.2) | *0.033* | 245 (29.8) | 33 (45.2) | 212 (28.3) | *0.003* |
| Femoral neck BMD ($g/cm^2$) | 0.829 ± 0.155 | 0.743 ± 0.132 | 0.843 ± 0.154 | *< 0.001* | 0.947 ± 0.141 | 0.888 ± 0.150 | 0.953 ± 0.139 | *0.001* |
| BMD status | | | | *< 0.001* | | | | *0.004* |
| Normal | 253 (31.3) | 13 (11.3) | 240 (34.6) | | 316 (38.5) | 19 (26.0) | 297 (39.7) | |
| Osteopenia | 399 (49.3) | 57 (49.6) | 342 (49.3) | | 454 (55.3) | 44 (60.3) | 410 (54.8) | |
| Osteoporosis | 157 (19.4) | 45 (39.1) | 112 (16.1) | | 51 (6.2) | 10 (13.7) | 41 (5.5) | |

Italic values indicate a significant difference between groups

† Too few to conduct statistical analysis

information about falls for calculation of Garvan scores. Five women were taking bisphosphonates and were also excluded, as these fracture risk calculators are intended for treatment naïve patients. This left 809 women to be included in the study. There were 920 men aged 50–90 years at baseline. Of these, 67 did not have BMD, 10 had insufficient information to calculate FRAX, six had weight over 125 kg (preventing calculation of FRAX scores) and 16 were taking bisphosphonates, leaving 821 available for inclusion into this study. During follow-up, 207 women (25.6%) and 240 men (29.2%) died.

Women ($n = 809$) were followed for a total of 7234 person-years and men ($n = 821$) for 7118 person-years. Over the follow-up period, 115 women sustained a MOF; 35 hip, 54 clinical spine, 21 proximal humerus, and 5 forearm fractures. In men, 73 sustained a MOF; 15 hip, 47 clinical spine, 2 proximal humerus, and 9 forearm. There were 184 fragility fractures according to Garvan criteria in women and 109 in men. Additional hip fractures following a MOF occurred in 7 women and 2 men, resulting in a total number of hip fractures of 42 and 17, respectively.

Table 1 lists the descriptive characteristics of the sample. Women and men with fracture were older, had lower femoral neck BMD, were more likely to have a prior fracture, use glucocorticoids, or have fallen over the previous 12 months than those who did not sustain a MOF. Women with MOF were also shorter and lighter than those that did not sustain a fracture. A similar proportion of women and men had normal BMD (T-score above −1.0) but women were less likely to have osteopenia (T-score −1.0 to −2.5) and more likely to have osteoporosis (T-score < −2.5).

Table 2 shows the number of fractures predicted by each FRAX and Garvan score and the number of fractures observed; $p > 0.05$ indicated that the calculators were effective in predicting fractures. Table 3 shows the results for predicted fractures divided into BMD categories. Table 4 shows the AUROC data and Table 5 shows the diagnostics indexes for each fracture risk assessment score.

## Predicting MOFs and fragility fractures in women

### FRAX

Neither $FRAX_{BMD}$ nor $FRAX_{noBMD}$ predicted MOFs well; underestimating the number of fractures by 54.8% and 46.1%, respectively. Both $FRAX_{BMD}$ and $FRAX_{noBMD}$ underestimated MOFs in the osteopenia and osteoporosis groups.

AUROCs for both $FRAX_{BMD}$ and $FRAX_{noBMD}$ were in the fair range (0.70–0.79). There was no significant difference between the AUROC values for $FRAX_{BMD}$ and $FRAX_{noBMD}$ when predicting MOFs in women (Supplement Fig. 1a). Although specificity was high (> 80%) for both scores, sensitivity was low (< 50%), indicating a poor model. Positive predictive values (PPVs) and negative predictive values (NPVs) were not different between $FRAX_{BMD}$ and $FRAX_{noBMD}$. $FRAX_{BMD}$ categorised fewer of the 115 women with MOFs to be at high risk for fracture (≥ 20%, $n = 27$) than $FRAX_{noBMD}$ ($n = 31$), though the difference was not statistically significant ($p = 0.543$).

### Garvan

Fragility fractures were underestimated by both $Garvan_{BMD}$ (24.5%) and $Garvan_{noBMD}$ (21.7%). Garvan scores accurately

**Table 2** The number of major osteoporotic (MOFs)/fragility and hip fractures predicted by FRAX and Garvan, compared against the number of fractures actually observed. Values in italics shows $p$ values ≤ 0.05, which indicates discordance between the predicted and observed number of fractures

| | Women | | | Men | | |
|---|---|---|---|---|---|---|
| | Predicted | Observed | $p$ | Predicted | Observed | $p$ |
| FRAX | | | | | | |
| MOF | | | | | | |
| $FRAX_{BMD}$ | 52 | 115 | < 0.001 | 26 | 73 | < 0.001 |
| $FRAX_{noBMD}$ | 62 | 115 | < 0.001 | 31 | 73 | < 0.001 |
| Hip | | | | | | |
| $FRAX_{BMD}$ | 20 | 42 | 0.004 | 10 | 17 | 0.174 |
| $FRAX_{noBMD}$ | 27 | 42 | 0.065 | 15 | 17 | 0.721 |
| Garvan | | | | | | |
| Fragility fractures | | | | | | |
| $Garvan_{BMD}$ | 139 | 184 | *0.005* | 88 | 109 | 0.109 |
| $Garvan_{noBMD}$* | 144 | 184 | *0.013* | 47 | 109 | *< 0.001* |
| Hip | | | | | | |
| $Garvan_{BMD}$ | 50 | 42 | 0.390 | 21 | 17 | 0.511 |
| $Garvan_{noBMD}$* | 12 | 42 | *< 0.001* | 29 | 17 | 0.073 |

*Garvan without bone mineral density (BMD) uses weight in the calculation instead

**Table 3** The number of major osteoporotic (MOFs)/fragility and hip fractures predicted by the FRAX and Garvan (with and without BMD measurements), compared against the number of fractures actually observed in BMD categories. Italicised *p* values indicate discordance between expected and observed values

| | Predicted | Observed | *p*† | Predicted | Observed | *p*† | Predicted | Observed | *p*† |
|---|---|---|---|---|---|---|---|---|---|
| Women | Normal ($n = 253$) | | | Osteopenia ($n = 399$) | | | Osteoporosis ($n = 157$) | | |
| FRAX | | | | | | | | | |
| MOF | | | | | | | | | |
| $FRAX_{BMD}$ | 7 | 13 | 0.171 | 24 | 57 | *< 0.001* | 21 | 45 | *0.001* |
| $FRAX_{noBMD}$ | 12 | 13 | 0.837 | 31 | 57 | *0.003* | 19 | 45 | *< 0.001* |
| Hip | | | | | | | | | |
| $FRAX_{BMD}$ | 1 | 4 | 0.178 | 8 | 19 | *0.031* | 11 | 19 | 0.125 |
| $FRAX_{noBMD}$ | 4 | 4 | 1.000 | 14 | 19 | 0.374 | 10 | 19 | 0.079 |
| Garvan | | | | | | | | | |
| Fragility fractures | | | | | | | | | |
| $Garvan_{BMD}$ | 20 | 29 | 0.176 | 70 | 90 | 0.077 | 50 | 65 | 0.079 |
| $Garvan_{noBMD}$* | 31 | 29 | 0.783 | 73 | 90 | 0.136 | 40 | 65 | *0.003* |
| Hip | | | | | | | | | |
| $Garvan_{BMD}$ | 2 | 4 | 0.411 | 19 | 19 | 1.000 | 29 | 19 | 0.117 |
| $Garvan_{noBMD}$* | 1 | 4 | 0.178 | 5 | 19 | *0.004* | 6 | 19 | *0.007* |
| Men | Normal ($n = 316$) | | | Osteopenia ($n = 454$) | | | Osteoporosis ($n = 51$) | | |
| FRAX | | | | | | | | | |
| MOF | | | | | | | | | |
| $FRAX_{BMD}$ | 6 | 19 | *0.008* | 16 | 44 | *< 0.001* | 3 | 10 | *0.038* |
| $FRAX_{noBMD}$ | 10 | 19 | 0.087 | 18 | 44 | *0.001* | 3 | 10 | *0.038* |
| Hip | | | | | | | | | |
| $FRAX_{BMD}$ | 1 | 4 | 0.178 | 6 | 10 | 0.313 | 2 | 3 | 0.647 |
| $FRAX_{noBMD}$ | 4 | 4 | 1.000 | 9 | 10 | 0.817 | 2 | 3 | 0.647 |
| Garvan | | | | | | | | | |
| Fragility fractures | | | | | | | | | |
| $Garvan_{BMD}$ | 19 | 29 | 0.133 | 56 | 62 | 0.554 | 12 | 18 | 0.192 |
| $Garvan_{noBMD}$* | 13 | 29 | *0.011* | 28 | 62 | *< 0.001* | 5 | 18 | *0.002* |
| Hip | | | | | | | | | |
| $Garvan_{BMD}$ | 1 | 4 | 0.178 | 12 | 10 | 0.666 | 8 | 3 | 0.110 |
| $Garvan_{noBMD}$* | 7 | 4 | 0.362 | 18 | 10 | 0.125 | 4 | 3 | 0.695 |

*Garvan without bone mineral density (BMD) uses weight in the calculation instead

† *p* value for comparison to normal BMD

predicted fragility fractures across the three BMD groups, except for $Garvan_{noBMD}$ in women with osteoporosis.

AUROC values were significantly different between $Garvan_{BMD}$ and $Garvan_{noBMD}$ (0.696, 95% CI 0.652–0.739 vs 0.668, 95% CI 0.623–0.712, $p = 0.008$), with $Garvan_{BMD}$ having a higher ROC area (Supplement Fig. 2a). However, both were in the poor range (< 0.70). Sensitivity was at least 75% for both $Garvan_{BMD}$ and $Garvan_{noBMD}$; however, specificity was lower. Only $Garvan_{BMD}$ had acceptable sensitivity and specificity, though specificity was only just above 50%. There were no differences between PPV or NPV values between $Garvan_{BMD}$ and $Garvan_{noBMD}$. $Garvan_{BMD}$ categorised more women with fragility fracture to be at high risk (≥ 14%, $n = 142$) compared to $Garvan_{noBMD}$ ($n = 139$); however, this difference was not statistically significant ($p = 0.713$).

## Predicting hip fractures in women

### FRAX

Although there appeared to be a difference between the predicted and observed number of fractures using $FRAX_{noBMD}$ (27 vs 42), this difference did not reach significance ($p = 0.065$). $FRAX_{BMD}$ underestimated the number of hip fractures

**Table 4** Receiver operator characteristic (ROC) areas for FRAX and Garvan prediction of major osteoporotic (MOF)/fragility or hip fracture in male and female participants. Data presented as mean (95% CI)

| | Women | Men |
|---|---|---|
| FRAX | | |
| MOF | | |
| $FRAX_{BMD}$ | 0.753 (0.709–0.796) | 0.723 (0.670–0.776) |
| $FRAX_{noBMD}$ | 0.738 (0.694–0.782) | 0.704 (0.649–0.760) |
| Hip | | |
| $FRAX_{BMD}$ | 0.775 (0.709–0.841) | 0.796 (0.709–0.883) |
| $FRAX_{noBMD}$ | 0.770 (0.703–0.837) | 0.758 (0.667–0.849) |
| Garvan | | |
| Fragility fractures | | |
| $Garvan_{BMD}$ | 0.696 (0.652–0.739) | 0.683 (0.631–0.734) |
| $Garvan_{noBMD}$* | 0.668 (0.623–0.712) | 0.667 (0.615–0.719) |
| Hip | | |
| $Garvan_{BMD}$ | 0.802 (0.734–0.869) | 0.799 (0.707–0.890) |
| $Garvan_{noBMD}$* | 0.792 (0.725–0.858) | 0.773 (0.691–0.855) |

*Garvan without bone mineral density (BMD) uses weight in the calculation instead

by 52.4%. When divided into BMD categories, $FRAX_{BMD}$ underestimated hip fractures in the osteopenia group.

AUROC values were in the fair range for both scores. There was no significant difference between the AUROC values for $FRAX_{noBMD}$ and $FRAX_{BMD}$ when predicting hip fractures in women (Supplement Fig. 1b). Sensitivity was higher for $FRAX_{noBMD}$ than $FRAX_{BMD}$, however specificity was lower, although both were in the acceptable range (sensitivity and specificity 50–80%). The PPV value appeared higher for $FRAX_{BMD}$ than $FRAX_{noBMD}$; however, this did not reach significance (11.5%, 95% CI 7.9–15.9 vs 9.4%, 95% CI 6.6–12.8). NPVs were similar between the two FRAX scores. Of the 42 women with hip fracture, $FRAX_{BMD}$ categorised fewer women with hip fracture to be at high risk ($n = 30$) than $FRAX_{noBMD}$ ($n = 34$), though the difference was not statistically significant ($p = 0.302$).

### Garvan

Hip fractures were underestimated by 71.4% for $Garvan_{noBMD}$; however, $Garvan_{BMD}$ was accurate, predicting 50 fractures (observed 42). $Garvan_{BMD}$ accurately predicted hip fractures across the three BMD groups, while $Garvan_{noBMD}$ underestimated hip fractures in women with osteopenia and osteoporosis.

The AUROC for $Garvan_{BMD}$ was good (0.80–0.89) and for $Garvan_{noBMD}$ it was fair (0.70–0.79). AUROC values were not significantly different between $Garvan_{BMD}$ and $Garvan_{noBMD}$ (Supplement Fig. 2b). Sensitivity and specificity for $Garvan_{BMD}$ was close to the acceptable range (specificity was only 49.7%). $Garvan_{noBMD}$ had acceptable sensitivity and specificity (between 50 and 80%); however, the sensitivity was only 59.5%, which may not be considered high enough for detection of hip fractures. PPV and NPV values were similar for the two scores. $Garvan_{BMD}$ categorised more women with hip fracture to be at high risk ($n = 39$) compared to $Garvan_{noBMD}$ ($n = 25$), and this difference was statistically significant ($p < 0.001$).

## Predicting MOFs and fragility fractures in men

### FRAX

Neither $FRAX_{BMD}$ nor $FRAX_{noBMD}$ predicted MOFs well, underestimating the number of fractures by 64.4% and 57.5%, respectively. When considering BMD categories, both scores underestimated the number of fractures in all groups, except for $FRAX_{noBMD}$ in the normal BMD group.

AUROC values for $FRAX_{BMD}$ and $FRAX_{noBMD}$ were both in the fair range. There was no significant difference between the AUROC values for $FRAX_{BMD}$ and $FRAX_{noBMD}$ (Supplement Fig. 1c). Despite high specificity, the sensitivity was low for both scores, thus resulting in poor prediction. The PPV value appeared to be higher for $FRAX_{BMD}$ than $FRAX_{noBMD}$; however, this did not reach significance (33.3%, 95% CI 0.8–90.6 vs 22.2%, 95% CI 2.8–60.0). NPV values were similar between the two scores. Of the 73 men with MOFs, only 1 and 2 were categorised as being at high risk by $FRAX_{BMD}$ and $FRAX_{noBMD}$, respectively. There was no difference between the scores ($p = 0.559$).

### Garvan

Fragility fractures were underestimated by 56.9% for $Garvan_{noBMD}$. $Garvan_{BMD}$, however, accurately predicted fragility fractures. The results were similar when stratified by BMD groups; $Garvan_{BMD}$ accurately predicted fragility fractures, while $Garvan_{noBMD}$ underestimated fractures across the three groups.

The AUROC values for $Garvan_{BMD}$ and $Garvan_{noBMD}$ were both in the poor range (< 0.70). AUROC values appeared to be different between $Garvan_{BMD}$ and $Garvan_{noBMD}$ (Supplement Fig. 2c); however, this did not quite reach significance ($p = 0.051$). Sensitivity was higher for $Garvan_{BMD}$ than $Garvan_{noBMD}$; however, specificity was lower. Overall $Garvan_{BMD}$ had acceptable sensitivity and specificity, while for $Garvan_{noBMD}$, it was poor. Both PPV and NPV were similar for both scores. Of the 109 men with fragility fracture, $Garvan_{BMD}$ categorised more to be at high risk (≥ 14%, $n = 67$) compared to $Garvan_{noBMD}$ ($n = 32$), and this difference was statistically significant ($p < 0.001$).

**Table 5** Diagnostics indexes (with 95% CI) for FRAX and Garvan prediction of major osteoporotic (MOF)/fragility or hip fracture in male and female participants. For FRAX, scores of ≥ 3% for hip and ≥ 20% for MOF were used to calculate the values. For Garvan, scores were ≥ 3% for hip and ≥ 14% for fragility fracture

| | Sensitivity | Specificity | Correctly classified | LR+ | LR- | PPV | NPV |
|---|---|---|---|---|---|---|---|
| Women | | | | | | | |
| FRAX | | | | | | | |
| MOF | | | | | | | |
| $FRAX_{BMD}$ | 23.5 (16.1–32.3) | 93.4 (91.3–95.1) | 83.4 | 3.54 (2.30–5.46) | 0.82 (0.74–0.91) | 37.0 (26.0–49.1) | 88.0 (85.5–90.3) |
| $FRAX_{noBMD}$ | 27.0 (19.1–36.0) | 89.9 (87.4–92.1) | 81.0 | 2.67 (1.84–3.88) | 0.81 (0.73–0.91) | 30.7 (21.9–40.7) | 88.1 (85.5–90.4) |
| Hip | | | | | | | |
| $FRAX_{BMD}$ | 71.4 (55.4–84.3) | 69.8 (66.4–73.0) | 69.8 | 2.36 (1.90–2.94) | 0.41 (0.25–0.66) | 11.5 (7.9–15.9) | 97.8 (96.2–98.9) |
| $FRAX_{noBMD}$ | 81.0 (65.9–91.4) | 57.1 (53.5–60.6) | 58.3 | 1.89 (1.60–2.23) | 0.33 (0.18–0.62) | 9.4 (6.6–12.8) | 98.2 (96.5–99.2) |
| Garvan | | | | | | | |
| Fragility fractures | | | | | | | |
| $Garvan_{BMD}$ | 77.2 (70.4–83.0) | 50.7 (46.7–54.7) | 56.7 | 1.57 (1.40–1.75) | 0.45 (0.34–0.59) | 31.6 (27.3–36.1) | 88.3 (84.5–91.4) |
| $Garvan_{noBMD}$* | 75.5 (68.7–81.6) | 45.6 (41.6–49.6) | 52.4 | 1.39 (1.25–1.55) | 0.54 (0.41–0.70) | 29.0 (25.0–33.3) | 86.4 (82.2–89.9) |
| Hip | | | | | | | |
| $Garvan_{BMD}$ | 92.9 (80.5–98.5) | 49.7 (46.1–53.3) | 51.9 | 1.85 (1.65–2.06) | 0.14 (0.05–0.43) | 9.2 (6.6–12.3) | 99.2 (97.7–99.8) |
| $Garvan_{noBMD}$* | 59.5 (43.3–74.4) | 85.8 (83.1–88.2) | 84.4 | 4.19 (3.09–5.68) | 0.47 (0.33–0.68) | 18.7 (12.5–26.3) | 97.5 (96.0–98.5) |
| Men | | | | | | | |
| FRAX | | | | | | | |
| MOF | | | | | | | |
| $FRAX_{BMD}$ | 1.4 (0.03–7.4) | 99.7 (99.0–100.0) | 91.0 | 5.12 (0.47–55.80) | 0.99 (0.96–1.02) | 33.3 (0.8–90.6) | 91.2 (89.0–93.0) |
| $FRAX_{noBMD}$ | 2.7 (0.3–9.6) | 99.1 (98.1–99.6) | 90.5 | 2.93 (0.62–13.80) | 0.98 (0.94–1.02) | 22.2 (2.8–60.0) | 91.3 (89.1–93.1) |
| Hip | | | | | | | |
| $FRAX_{BMD}$ | 41.2 (18.4–67.1) | 85.9 (83.3–88.3) | 85.0 | 2.93 (1.62–5.30) | 0.68 (0.46–1.02) | 5.8 (2.4–11.6) | 98.6 (97.4–99.3) |
| $FRAX_{noBMD}$ | 76.5 (50.1–93.2) | 69.2 (65.8–72.3) | 69.3 | 2.48 (1.87–3.29) | 0.34 (0.14–0.80) | 5.0 (2.7–8.4) | 99.3 (98.2–99.8) |
| Garvan | | | | | | | |
| Fragility fractures | | | | | | | |
| $Garvan_{BMD}$ | 61.5 (51.7–70.6) | 66.2 (62.5–69.6) | 65.5 | 1.82 (1.52–2.18) | 0.58 (0.46–0.74) | 21.8 (17.3–26.8) | 91.8 (89.1–94.0) |
| $Garvan_{noBMD}$* | 29.4 (21.0–38.8) | 82.7 (79.7–85.4) | 75.6 | 1.70 (1.22–2.37) | 0.85 (0.75–0.97) | 20.6 (14.6–27.9) | 88.4 (85.8–90.8) |
| Hip | | | | | | | |
| $Garvan_{BMD}$ | 76.5 (50.1–93.2) | 70.5 (67.2–73.7) | 70.7 | 2.59 (1.95–3.45) | 0.33 (0.14–0.79) | 5.2 (2.8–8.7) | 99.3 (98.2–99.8) |
| $Garvan_{noBMD}$* | 88.2 (63.6–98.5) | 56.8 (53.3–60.3) | 57.5 | 2.04 (1.69–2.47) | 0.21 (0.06–0.76) | 4.1 (2.3–6.7) | 99.6 (98.4–99.9) |

*Garvan without bone mineral density (BMD) uses weight in the calculation instead

## Predicting hip fractures in men

### FRAX

Both $FRAX_{BMD}$ and $FRAX_{noBMD}$ predicted the number of hip fractures well. This was also observed when the analysis was completed for the BMD categories.

AUROC values were in the fair range for both scores. There was no significant difference between the AUROC values (Supplement Fig. 1d). $FRAX_{noBMD}$ had acceptable sensitivity and specificity; however, for $FRAX_{BMD}$, it was poor. PPV and NPV values were similar for $FRAX_{BMD}$ and $FRAX_{noBMD}$, Of the 17 men with hip fracture, 7 and 13 were considered to be at high risk by $FRAX_{BMD}$ and $FRAX_{noBMD}$, respectively, and the difference was statistically significant ($p = 0.025$).

### Garvan

Hip fractures were accurately predicted by both $Garvan_{BMD}$ and $Garvan_{noBMD}$. This was also observed when the analysis was stratified by BMD groups.

AUROC values were both in the fair range; however, for $Garvan_{BMD}$, it was very close to good (0.799). AUROC values were not significantly different between $Garvan_{BMD}$ and $Garvan_{noBMD}$ (Supplement Fig. 2d). Both $Garvan_{BMD}$ and $Garvan_{noBMD}$ had acceptable sensitivity and specificity

(both between 50 and 80%). PPV and NPV values were similar for both $Garvan_{BMD}$ and $Garvan_{noBMD}$. $Garvan_{BMD}$ categorised fewer of the 17 men with hip fracture to be at high risk ($n = 13$) compared to $Garvan_{noBMD}$ ($n = 15$), but this difference was not statistically significant ($p = 0.362$).

## Discussion

This study examined the discriminative ability of the Australian FRAX calculator and the Garvan calculator in a population-based Australian sample. Overall, FRAX underestimated MOFs in both men and women, regardless of whether BMD was added to the calculation. FRAX accurately predicted hip fractures, except for $FRAX_{BMD}$ in women. The Garvan calculator underestimated fragility fractures, except for $Garvan_{BMD}$ in men. Only $Garvan_{noBMD}$ in women was not accurate in predicting hip fractures. Both MOF and hip fractures were underestimated in the osteopenia and osteoporosis BMD groups. Only MOFs by $FRAX_{BMD}$ and fragility fractures by $Garvan_{noBMD}$, both in men, were underestimated in the normal BMD group. AUROCs were not significantly different between scores with and without BMD, except for Garvan predicting fragility fractures in both men and women.

According to sensitivity and specificity comparisons, overall FRAX scores performed poorly for predicting MOFs in both men and women, but performed acceptably well for predicting hip fractures (except for $FRAX_{BMD}$ in men). $Garvan_{BMD}$ had acceptable sensitivity and specificity for predicting fragility fractures in men and women, but had poor AUROC values. $Garvan_{noBMD}$ however, performed poorly in predicting fragility fractures in both sexes. For hip fracture predictions, $Garvan_{BMD}$ and $Garvan_{noBMD}$ both performed acceptably well and with BMD, had values in the "good" range for AUROCs.

A possible explanation of the relative underestimation of fractures by FRAX could be its design, which adjusts fracture risk based on expected mortality rates, resulting in a decline of absolute risk for those aged over 80 years [18]. This would result in a lower calculated fracture risk in the elderly population than would be predicted based on BMD and risk factors independent of expected mortality. It is possible that individuals who might benefit from anti-fracture therapies might not be treated if their eligibility for treatment were based on the Australian FRAX tool.

Where data regarding the rates of major osteoporotic and hip fractures are unavailable, the Australian FRAX calculator assumes that the age- and sex- specific pattern of these fractures is similar to that observed in Sweden [30]. Our study suggests that this assumption may not be correct. Hip fracture risks may also be underestimated by the FRAX calculator due to other limitations including some clinical risk factors missing (e.g., falls, which are not used in the FRAX calculator [31]) and binary responses to clinical risk factors that likely have a dose-effect (e.g., glucocorticoids, smoking) [30]. Additionally, the prevalence of clinical risk factors are different between the Dubbo study [12], which was used to develop the Australian FRAX calculator, and this study, which may be leading to poor calibration of fracture risk assessment.

Another potential reason for differences in the number of predicted and observed number of fractures could be variation in fracture rates. Indeed, among the Australian population-based studies, the Dubbo Osteoporosis Epidemiology Study has the highest point estimate for hip fracture [32] and would be expected to therefore generate the highest numbers of MOFs, whereas our analysis suggests that the application of this assumption underestimates fracture numbers. It has also been reported that fracture rates vary by region, even within the same country [30, 32]. Where possible, FRAX has used national rather than regional data, however these data are not available for Australia. This could be contributing to the underestimation of fractures, and highlights the need for collection of national data to update the Australian FRAX calculator. Additionally, in this study, we used the US National Osteoporosis Foundation cut-points to define high risk for MOF and hip fractures using FRAX. These cut-points may not be appropriate for Australian populations, and further research should be completed to determine the optimal cut-points for use in Australia.

Only a few studies have used both the FRAX and Garvan calculators to assess fracture risk. As far as we know, no studies have compared the two in an Australian cohort, the population in which the Garvan calculator was derived. A New Zealand study compared the number of hip, osteoporotic and fragility fractures predicted by the FRAX (New Zealand) and Garvan calculators in 1422 women followed for a mean of 8.8 years (range 0.2–11.4 years) [15]. Inconsistent with what is reported in our study, the study reported that $FRAX_{BMD}$ predicted hip fractures well, whereas both $FRAX_{noBMD}$ and the $Garvan_{BMD}$ calculator overestimated hip fracture numbers. The study also reported that $FRAX_{BMD}$ and $FRAX_{noBMD}$ both underestimated the number of osteoporotic fracture whereas $Garvan_{BMD}$ effectively predicted the risk of fragility fractures. Another study found FRAX to underestimate the number of osteoporotic fractures within a Polish cohort of 501 women, both with and without BMD measurements [17]. The study used FRAX (UK) and was unable to determine the dates when fractures occurred, raising the possibility that fractures observed in the study happened beyond the 10-year period. A recently published study by Crandall et al. [33] showed that the US FRAX and Garvan calculators performed poorly, with low sensitivity, specificity, and AUROCs, for discriminating fractures in postmenopausal women aged 50–64 years. Studies have traditionally compared predictive calculators using AUROC curves [14, 15, 34]. The current study is consistent with previous studies [15, 34] that showed similar

AUROC curves for all scores. Consistent with all previous studies, the FRAX calculator showed discriminatory abilities that were better than chance.

Although fractures were underestimated in this study by both FRAX and Garvan, AUROCs showed that the calculators performed well in our sample, when compared to several other previously published studies. For example, Bolland et al. [15] reported AUROC values ranging from 0.60 to 0.70 using the New Zealand FRAX tool and 0.63 and 0.67 using the Garvan algorithm. Additionally, González-Macías et al. [20] reported AUROCs of 0.615 for MOFs and 0.640 for hip fracture prediction using the Spanish FRAX tool in women aged ≥ 65 years from Spain. AUROCs in our study were also higher compared to those reported by Crandall et al. [33]; the AUROC values for FRAX ranged from 0.58 to 0.68 for FRAX and 0.55 to 0.62 for Garvan.

Strengths of this study include radiologically confirmed fractures, which allowed determination of fracture endpoints, even if participants did not participate in subsequent follow-up visits. This study also includes both men and women, and data on parental hip fracture were available, which is often missing from studies involving FRAX. Limitations include assuming uniform risk for fracture over each year during the period participants were followed. Since we censored individuals at date of death, the 10-year observed fracture probability will be overestimated compared to FRAX predictions and prevent comparison with Garvan scores, as competing mortality is handled differently by the two calculators. Small numbers of fractures, particularly for men and hip fractures, was another limitation; however, it has been argued that smaller sample sizes may be viable as long as it is representative of the population in question [18]. It is also possible that some fractures were missed if no X-ray was performed or if a participant sustained a fracture and were managed outside the study region. The participants in this study were a randomly selected sample recruited from a population that resembles the Australian white population; however, the population of Australia also has a range of mixed ethnicities which varies across different regions. We used the Australian FRAX calculator and thus results cannot be extrapolated to FRAX calculators for other countries. The Garvan fracture risk calculator is also only validated for Australian populations. Finally, information including prior fractures, parental hip fractures, glucocorticoid usage, secondary osteoporosis, alcohol consumption, and falls were obtained by self-report, which may involve recall bias. It has been shown, however, that 12-month recall of falls accurately predicts falls collected prospectively [35].

## Conclusion

In this cohort of men and women, the Australian FRAX calculator substantially underestimated the number of MOFs and the Garvan calculator underestimated the number of fragility fractures. The underestimation of MOFs by FRAX was greater than the underestimation of fragility fractures by Garvan. Hip fractures were predicted well by both calculators in both men and women with a few exceptions. Underestimation of fractures using both scores occurred primarily in individuals with osteopenia and osteoporosis. There were no significant differences in the AUROC curves, with the exception of Garvan predicting fragility fractures, though there were differences in sensitivity and specificity, demonstrating that scores may predict different number of fractures but still have similar AUROCs. Sensitivity and specificity analyses indicated that MOF and fragility fractures were poorly predicted; however, the FRAX and Garvan calculators performed acceptably well for hip fractures. Overall, hip fractures were predicted acceptably well by both calculators; however, MOFs were underestimated by FRAX more than fragility fractures by Garvan.

**Acknowledgments** We thank the Australian Institute for Health and Welfare (AIHW) for providing deaths data.

**Funding information** This study is supported by the National Health and Medical Research Council (NHMRC) of Australia (251638, 299831, 628582), the Geelong Regional Medical Foundation, Arthritis Foundation of Australia, and Perpetual Trustees.

## Compliance with ethical standards

**Conflict of interest** KLH-K was supported by an Alfred Deakin Postdoctoral Research Fellowship. KBA was supported by an Australian Government Research Training Program Scholarship. JAP, NAP and MAK have received speaker fees from Amgen. GCN received speaker fees from Amgen, Novartis, MSD, Eli Lilly and Sanofi-Aventis. None is related to this manuscript. All other authors state that they have no conflict of interest.

## References

1. Cooper C (1993) Epidemiology and public health impact of osteoporosis. Baillieres Clin Rheumatol 7:459–477. https://doi.org/10.1016/S0950-3579(05)80073-1
2. Pasco JA, Henry MJ, Korn S, Nicholson GC, Kotowicz MA (2009) Morphometric vertebral fractures of the lower thoracic and lumbar spine, physical function and quality of life in men. Osteoporos Int 20:787–792. https://doi.org/10.1007/s00198-008-0744-6
3. Otmar R, Kotowicz MA, Brennan SL, Bucki-Smith G, Korn S, Pasco JA (2013) Personal and psychosocial impacts of clinical fracture in men. J Mens Health 10:22–27
4. Pasco JA, Sanders KM, Hoekstra FM, Henry MJ, Nicholson GC, Kotowicz MA (2005) The human cost of fracture. Osteoporos Int 16:2046–2052
5. Cummings SR, Melton LJ (2002) Epidemiology and outcomes of osteoporotic fractures. Lancet 359:1761–1767. https://doi.org/10.1016/S0140-6736(02)08657-9
6. Sanders KM, Nicholson GC, Watts JJ, Pasco JA, Henry MJ, Kotowicz MA, Seeman E (2006) Half the burden of fragility

fractures in the community occur in women without osteoporosis. When is fracture prevention cost-effective? Bone 38:694–700
7. Reid IR, Horne AM, Mihov B, Stewart A, Garratt E, Wong S, Wiessing KR, Bolland MJ, Bastin S, Gamble GD (2018) Fracture prevention with Zoledronate in older women with osteopenia. N Engl J Med 379:2407–2416. https://doi.org/10.1056/NEJMoa1808082
8. Kanis JA (2008) FRAX™ and the assessment of fracture probability in men and women from the UK. Osteoporos Int 19:385–397
9. University of Sheffield UK (2011) FRAX ® WHO Fracture Risk Assessment Tool. http://www.shef.ac.uk/FRAX/
10. Bauer DC (2011) FRAX, falls, and fracture prediction: predicting the future. Arch Intern Med 171:1661–1662
11. Cosman F, de Beur SJ, LeBoff MS, Lewiecki EM, Tanner B, Randall S, Lindsay R, National Osteoporosis Foundation (2014) Clinician's guide to prevention and treatment of osteoporosis. Osteoporos Int 25:2359–2381. https://doi.org/10.1007/s00198-014-2794-2
12. Nguyen ND, Frost SA, Center JR et al (2008) Development of prognostic nomograms for individualizing 5-year and 10-year fracture risks. Osteoporos Int 19:1431–1444. https://doi.org/10.1007/s00198-008-0588-0
13. Nguyen ND, Frost SA, Center JR, et al (2007) Development of a nomogram for individualizing hip fracture risk in men and women. Osteoporos Int 18:1109–1117. https://doi.org/10.1007/s00198-007-0362-8
14. Henry MJ, Pasco JA, Merriman EN, Zhang Y, Sanders KM, Kotowicz MA, Nicholson GC (2011) Fracture risk score and absolute risk of fracture. Radiology 259:495–501. https://doi.org/10.1148/radiol.10101406
15. Bolland MJ, Siu ATY, Mason BH, Horne AM, Ames RW, Grey AB, Gamble GD, Reid IR (2011) Evaluation of the FRAX and Garvan fracture risk calculators in older women. J Bone Miner Res 26:420–427. https://doi.org/10.1002/jbmr.215
16. Sandhu S, Nguyen N, Center J, et al (2010) Prognosis of fracture: evaluation of predictive accuracy of the FRAX algorithm and Garvan nomogram. Osteoporos Int 21:863–871. https://doi.org/10.1007/s00198-010-1527-4
17. Czerwinski E, Kanis JA, Osieleniec J, Kumorek A, Milert A, Johansson H, McCloskey EV, Gorkiewicz M (2011) Evaluation of FRAX to characterise fracture risk in Poland. Osteoporos Int 22:2507–2512. https://doi.org/10.1007/s00198-010-1502-0
18. Kanis JA, Oden A, Johansson H, McCloskey E (2012) Pitfalls in the external validation of FRAX. Osteoporos Int 23:423–431. https://doi.org/10.1007/s00198-011-1846-0
19. Azagra R, Roca G, Encabo G, Aguyé A, Zwart M, Güell S, Puchol N, Gene E, Casado E, Sancho P, Solà S, Torán P, Iglesias M, Gisbert MC, López-Expósito F, Pujol-Salud J, Fernandez-Hermida Y, Puente A, Rosàs M, Bou V, Antón JJ, Lansdberg G, Martín-Sánchez JC, Díez-Pérez A, Prieto-Alhambra D (2012) FRAX® 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 13:204. https://doi.org/10.1186/1471-2474-13-204
20. González-Macías J, Marin F, Vila J, Díez-Pérez A (2012) Probability of fractures predicted by FRAX® and observed incidence in the Spanish ECOSAP study cohort. Bone 50:373–377. https://doi.org/10.1016/j.bone.2011.11.006
21. Pasco JA, Nicholson GC, Kotowicz MA (2012) Cohort profile: Geelong osteoporosis study. Int J Epidemiol 41:1565–1575
22. Pasco JA, Nicholson GC, Henry MJ et al (1999) Identification of incident fractures: the Geelong osteoporosis study. Aust NZ J Med 29:203–206
23. Genant HK, Wu CY, van Kuijk C, Nevitt MC (1993) Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res 8:1137–1148. https://doi.org/10.1002/jbmr.5650080915
24. Giles GG, Ireland PD (1996) Dietary questionnaire for epidemiological studies (version 2). Melbourne, Cancer Counc Victoria
25. Henry MJ, Pasco JA, Korn S, Gibson JE, Kotowicz MA, Nicholson GC (2010) Bone mineral density reference ranges for Australian men: Geelong osteoporosis study. Osteoporos Int 21:909–917
26. Henry MJ, Pasco JA, Pocock NA, Nicholson GC, Kotowicz MA (2004) Reference ranges for bone densitometers adopted Australia-wide: Geelong osteoporosis study. Australas Radiol 48:473–475
27. Swets JA (1988) Measuring the accuracy of diagnostic systems. Science 240:1285–1293
28. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845
29. Garvan Institute of Medical Research (2008) Bone Fracture Risk Calculator. https://www.garvan.org.au/promotions/bone-fracture-risk/calculator/index.php. Accessed 8 May 2019
30. Kanis JA, Hans D, Cooper C et al (2011) Interpretation and use of FRAX in clinical practice. Osteoporos Int 22:2395–2411
31. Holloway KL, Kotowicz MA, Lane SE, Brennan SL, Pasco JA (2015) FRAX (Aus) and falls risk: association in men and women. Bone 76:1–4. https://doi.org/10.1016/j.bone.2015.03.004
32. Cooley H, Jones G (2001) A population-based study of fracture incidence in southern Tasmania: lifetime fracture risk and evidence for geographic variations within the same country. Osteoporos Int 12:124–130
33. Crandall CJ, Larson J, LaCroix A, Cauley JA, LeBoff MS, Li W, LeBlanc ES, Edwards BJ, Manson JAE, Ensrud K (2019) Predicting fracture risk in younger postmenopausal women: comparison of the Garvan and FRAX risk calculators in the Women's health initiative study. J Gen Intern Med 34:235–242
34. 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 25:2101–2107. https://doi.org/10.1002/jbmr.106
35. Sanders KM, Stuart AL, Scott D, Kotowicz MA, Nicholson GC (2015) Validity of 12-month falls recall in community-dwelling older women participating in a clinical trial. Int J Endocrinol 2015:210527. https://doi.org/10.1155/2015/210527