



**University of Dundee**

**Derivation and validation of the CFracture competing risk fracture prediction tool compared with QFracture in older people and people with comorbidity**

Livingstone, Shona J.; Guthrie, Bruce; McMinn, Megan; Eke, Chima; Donnan, Peter T.; Morales, Daniel R.

*Published in:*  
The Lancet Healthy Longevity

*DOI:*  
[10.1016/S2666-7568\(22\)00290-2](https://doi.org/10.1016/S2666-7568(22)00290-2)

*Publication date:*  
2023

*Licence:*  
CC BY

*Document Version*  
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

*Citation for published version (APA):*

Livingstone, S. J., Guthrie, B., McMinn, M., Eke, C., Donnan, P. T., & Morales, D. R. (2023). Derivation and validation of the CFracture competing risk fracture prediction tool compared with QFracture in older people and people with comorbidity: a population cohort study. *The Lancet Healthy Longevity*, 4(1), e43-e53. [https://doi.org/10.1016/S2666-7568\(22\)00290-2](https://doi.org/10.1016/S2666-7568(22)00290-2)

**General rights**

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Derivation and validation of the CFracture competing risk fracture prediction tool compared with QFracture in older people and people with comorbidity: a population cohort study



Shona J Livingstone, Bruce Guthrie, Megan McMinn, Chima Eke, Peter T Donnan, Daniel R Morales



## Summary

**Background** UK guidelines recommend the QFracture tool to predict the risk of major osteoporotic fracture and hip fracture, but QFracture calibration is poor, partly because it does not account for competing mortality risk. The aim of this study was to derive and validate a competing risk model to predict major osteoporotic fracture and hip fracture (CFracture) and compare its performance with that of QFracture in UK primary care.

**Methods** We used UK linked primary care data from the Clinical Practice Research Datalink GOLD database to identify people aged 30–99 years, split into derivation and validation cohorts. In the derivation cohort, we derived models (CFracture) using the same covariates as QFracture with Fine-Gray competing risk modelling, and included the Charlson Comorbidity Index score as an additional predictor of non-fracture death. In a separate validation cohort, we examined discrimination (using Harrell's C-statistic) and calibration of CFracture compared with QFracture. Reclassification analysis examined differences in the characteristics of patients reclassified as higher risk by CFracture but not by QFracture.

**Findings** The derivation cohort included 1 831 606 women and 1 789 820 men, and the validation cohort included 915 803 women and 894 910 men. Overall discrimination of CFracture was excellent (C-statistic=0.813 [95% CI 0.810–0.816] for major osteoporotic fracture and 0.914 [0.908–0.919] for hip fracture in women; 0.734 [0.729–0.740] for major osteoporotic fracture and 0.886 [0.877–0.895] for hip fracture in men) and was similar to QFracture. CFracture calibration overall and in people younger than 75 years was generally excellent. CFracture overpredicted major osteoporotic fracture and hip fracture in older people and people with comorbidity, but was better calibrated than QFracture. Patients classified as high-risk by CFracture but not by QFracture had a higher prevalence of current smoking and previous fracture, but lower prevalence of dementia, cancer, cardiovascular disease, renal disease, and diabetes.

**Interpretation** CFracture has similar discrimination to QFracture but is better calibrated overall and in younger people. Both models performed poorly in adults aged 85 years and older. Competing risk models should be recommended for fracture risk prediction to guide treatment recommendations.

**Funding** National Institute for Health and Care Research, Wellcome Trust, Health Data Research UK.

**Copyright** © 2023 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

## Introduction

Fragility fractures are a major cause of morbidity and mortality. The disease burden associated with fragility fractures has increased globally and is largest in India, China, the USA, Japan, and Germany.<sup>1</sup> In England and Wales, over 2 million people have osteoporosis, with one in three women and one in five men estimated to experience a fracture during their lifetime.<sup>2</sup> Osteoporosis treatment options include antiresorptive agents (eg, bisphosphonates, denosumab, raloxifene, and hormone replacement therapy) and anabolic agents (eg, teriparatide and romosozumab). These treatments are often recommended by international guidelines on the basis of clinical, bone density, or risk prediction stratification. In the UK, risk prediction models are recommended to

calculate the 10-year risk of major osteoporotic fractures (ie, wrist, proximal humerus, vertebral, or hip fractures) or hip fracture alone to guide decisions about investigation of bone mineral density and initiation of preventive treatment.<sup>2</sup> The QFracture and FRAX<sup>3,4</sup> tools are recommended by the UK National Institute for Health and Care Excellence (NICE) for risk stratification in all women older than 65 years and men older than 75 years, or in individuals older than 50 years with additional risk factors. For example, measurement of bone mineral density is recommended for people with a 10-year fracture risk of 10% or greater, although local pathways might vary. In contrast, the UK National Osteoporosis Guideline Group recommends FRAX, and outside the UK other models such as the Garvan Fracture Risk Calculator are

*Lancet Healthy Longev* 2023; 4: e43–53

See [Comment](#) page e6

Division of Population Health and Genomics, University of Dundee, Dundee, UK (S J Livingstone MSc, Prof P T Donnan PhD, D R Morales PhD); Department of Public Health, University of Southern Denmark, Odense, Denmark (D R Morales); Advanced Care Research Centre, Usher Institute, University of Edinburgh, Edinburgh, UK (Prof B Guthrie PhD, M McMinn PhD, C Eke PhD)

Correspondence to: Dr Daniel R Morales, Division of Population Health and Genomics, University of Dundee, Dundee DD2 4BF, UK [d.r.z.morales@dundee.ac.uk](mailto:d.r.z.morales@dundee.ac.uk)

### Research in context

#### Evidence before this study

Decisions to start long-term medication to prevent fracture events are guided by estimation of fracture risk, with investigation or treatment offered if patients exceed a particular risk threshold. UK National Institute of Health and Care Excellence guidelines recommend the QFracture risk prediction tool to inform clinical decisions, but QFracture has been shown to underpredict fracture risk, particularly in young people, and to overpredict fracture risk in older adults and individuals with multimorbidity. We searched PubMed from inception to June 16, 2022, for articles in English, using the search strategy [fracture[Title/Abstract] AND (predict\*[Title/Abstract]) AND ('competing risk'[Title/Abstract]) AND (osteoporosis[Title/Abstract]) AND (mortality[Title/Abstract])]. We identified seven articles, of which four performed a competing risk analysis in the context of fracture prediction. All studies predicted fracture risk in subpopulations only. Two studies were conducted in older people ( $\geq 60$  years) to quantify the effects of the predictors on the transition risks to fracture and mortality over 5 years, whereas the other quantified residual lifetime fracture risk. One study estimated

fracture risk in people with type 2 diabetes, finding that failing to account for competing risk mortality overestimates fracture risk. Lastly, a competing risk model was used to predict fracture risk in post-menopausal women only, showing good discrimination and calibration.

#### Added value of this study

This study shows that accounting for competing mortality risk and modelling a predictor of mortality leads to more accurate (better calibrated) prediction of major osteoporotic fracture and hip fracture risk than QFracture. This new model, CFracture, also reclassifies patients with different characteristics, leading to a lower estimated number needed to treat to prevent a fracture, depending on different levels of predicted risk and fracture type.

#### Implications of all the available evidence

CFracture has better calibration than QFracture and might be more accurate in older people and in those with multimorbidity. Prediction might be improved using CFracture or an alternative competing risk model instead of QFracture. Prediction models aimed at older people and those with multimorbidity should consider accounting for competing risk.

used.<sup>5,6</sup> Unlike FRAX, QFracture does not account for competing mortality risk, but in contrast with QFracture, the FRAX prediction equation has never been published.

Cox models are commonly used to estimate how predictors affect the hazard of an outcome in survival analysis. Censored data (due to loss to follow-up) are common in such analyses. Cox models and the Kaplan-Meier estimate of risk assume that individuals lost to follow-up have the same fracture risk (or other predicted outcome) as those who remain in follow-up. This assumption is false for those who die before a fracture event, which is more likely in older people and those with comorbidity.<sup>7-9</sup> These types of events (ie, deaths from other causes) are referred to as competing risk events. Failure to account for competing risk events leads to systematic overprediction of risk using standard Cox regression models, although the effect of this overprediction depends on how frequently competing risk events occur.<sup>10</sup>

Prediction tools that do not account for competing mortality risk will overestimate fracture risk in older people and those with multimorbidity, meaning that some patients might be unnecessarily recommended for treatments that come with a risk of harm and treatment disutility. Despite having excellent discrimination in the population as a whole, QFracture has recently been shown to systematically overpredict fracture risk in older and comorbid people with competing mortality risks, while simultaneously underpredicting fracture risk in younger and healthier people due to insufficient ascertainment of fracture events in the original derivation study.<sup>11</sup>

The aim of this study was to derive and internally validate a prediction tool for major osteoporotic fracture

and hip fracture that accounts for competing mortality risk, and to compare the new model's performance with that of QFracture.

## Methods

### Data source and population

We performed a cohort study using data from patients in the Clinical Practice Research Datalink (CPRD) GOLD database.<sup>12,13</sup> CPRD-GOLD contains electronic health records from primary care, including data on health conditions, prescriptions, laboratory measurements (taken in primary care), and lifestyle values, with linked data for UK hospitalisation and death registration, and is broadly representative of the UK population. Eligible patients had to be permanently registered with a general practice contributing up-to-standard data for at least 1 year; have linkage to Hospital Episodes Statistics and Office for National Statistics mortality data; be aged 30–99 years; and have observable records on or after Jan 1, 2004. Cohort exit was the earliest of: first fracture outcome event; non-fracture death; deregistration from the general practice; date of the last data collection from the practice; or March 31, 2016. The study was approved by the Medicines and Healthcare products Regulatory Agency's Independent Scientific Advisory Committee for database studies (reference ISAC 16/248) and was therefore exempt from ethical approval and patient consent.

### Outcomes

The outcomes of interest were major osteoporotic fracture and hip fracture, as in QFracture.<sup>3</sup> Major osteoporotic fracture was defined as hip, vertebral, wrist, or proximal

	Women		Men	
	Derivation cohort (n=1831606)	Internal validation cohort (n=915803)	Derivation cohort (n=1789820)	Internal validation cohort (n=894910)
Age, years	47.0 (35.0–63.0)	47.0 (35.0–63.0)	45.0 (35.0–59.0)	45.0 (35.0–59.0)
BMI, kg/m <sup>2</sup>	25.4 (22.4–29.6)	25.4 (22.4–29.6)	26.5 (23.9–29.6)	26.5 (23.9–29.5)
Ethnicity				
White or not recorded	1742593 (95.1%)	871830 (95.2%)	1704753 (95.2%)	852170 (95.2%)
Black African	15141 (0.8%)	7595 (0.8%)	13769 (0.8%)	7007 (0.8%)
Black Caribbean	3147 (0.2%)	1633 (0.2%)	2698 (0.2%)	1332 (0.1%)
Indian	17065 (0.9%)	8355 (0.9%)	18025 (1.0%)	9062 (1.0%)
Pakistani	7402 (0.4%)	3719 (0.4%)	8244 (0.5%)	4072 (0.5%)
Bangladeshi	2386 (0.1%)	1087 (0.1%)	3309 (0.2%)	1663 (0.2%)
Chinese	4960 (0.3%)	2398 (0.3%)	3693 (0.2%)	1824 (0.2%)
Other Asian ethnic group	12594 (0.7%)	6302 (0.7%)	11912 (0.7%)	5846 (0.7%)
Other ethnic group	26318 (1.4%)	12884 (1.4%)	23417 (1.3%)	11934 (1.3%)
Smoking status				
Non-smoker	989862 (64.4%)	494879 (64.4%)	710319 (53.8%)	355153 (53.8%)
Former smoker	260392 (16.9%)	130128 (16.9%)	293170 (22.2%)	146333 (22.2%)
Light smoker (<10 cigarettes per day)	89895 (5.8%)	45377 (5.9%)	83545 (6.3%)	41684 (6.3%)
Moderate smoker (10–19 cigarettes per day)	125373 (8.2%)	62705 (8.2%)	127193 (9.6%)	63797 (9.7%)
Heavy smoker (≥10 cigarettes per day)	71659 (4.7%)	35629 (4.6%)	105274 (8.0%)	52860 (8.0%)
Not recorded*	294425 (16.1%)	147085 (16.1%)	470319 (26.3%)	235083 (26.3%)
Alcohol consumption status				
None	380866 (27.9%)	190034 (27.8%)	210979 (17.4%)	106229 (17.5%)
<1 unit per day	569994 (41.7%)	284482 (41.7%)	366143 (30.2%)	182618 (30.2%)
1–2 units per day	373825 (27.4%)	187778 (27.5%)	447029 (36.9%)	222747 (36.8%)
3–6 units per day	35237 (2.6%)	17548 (2.6%)	149795 (12.4%)	74712 (12.3%)
7–9 units per day	3777 (0.3%)	1973 (0.3%)	25557 (2.1%)	12716 (2.1%)
>9 units per day	2046 (0.1%)	947 (0.1%)	12954 (1.1%)	6629 (1.1%)
Not recorded*	465861 (25.4%)	233041 (25.4%)	577363 (32.3%)	289259 (32.3%)
Previous major osteoporotic fracture	100490 (5.5%)	50095 (5.5%)	75451 (4.2%)	37801 (4.2%)
Parental history of osteoporosis or hip fracture	6940 (0.4%)	3621 (0.4%)	730 (0.1%)	347 (0.1%)
Nursing or care home resident	11165 (0.6%)	5654 (0.6%)	4994 (0.3%)	2461 (0.3%)
Type 1 diabetes	5809 (0.3%)	2938 (0.3%)	8117 (0.5%)	3891 (0.4%)
Type 2 diabetes	54513 (3.0%)	27202 (3.0%)	66585 (3.7%)	33424 (3.7%)
History of falls	102627 (5.6%)	51214 (5.6%)	49372 (2.8%)	24996 (2.8%)
Dementia	23257 (1.3%)	11635 (1.3%)	10049 (0.6%)	4987 (0.6%)
Cancer	62604 (3.4%)	31486 (3.4%)	45104 (2.5%)	22276 (2.5%)
Asthma or chronic obstructive pulmonary disease	236545 (12.9%)	118469 (12.9%)	202271 (11.3%)	101270 (11.3%)
Cardiovascular disease	104111 (5.7%)	52466 (5.7%)	130431 (7.3%)	64947 (7.3%)
Chronic liver disease	4072 (0.2%)	2021 (0.2%)	4461 (0.2%)	2292 (0.3%)
Chronic renal disease	22095 (1.2%)	11179 (1.2%)	16189 (0.9%)	8206 (0.9%)
Parkinson's disease	5061 (0.3%)	2524 (0.3%)	5649 (0.3%)	2699 (0.3%)
Rheumatoid arthritis or systemic lupus erythematosus	22037 (1.2%)	10913 (1.2%)	7984 (0.4%)	3986 (0.4%)
Malabsorption	23197 (1.3%)	11687 (1.3%)	18141 (1.0%)	8981 (1.0%)
Endocrine disorders	16792 (0.9%)	8297 (0.9%)	3929 (0.2%)	1937 (0.2%)
Epilepsy or prescribed anticonvulsants	43929 (2.4%)	22216 (2.4%)	39554 (2.2%)	19660 (2.2%)
Prescribed antidepressants	199831 (10.9%)	99744 (10.9)	95741 (5.3%)	47328 (5.3%)
Prescribed corticosteroids	24791 (1.4%)	12378 (1.4%)	15028 (0.8%)	7604 (0.8%)

Data are median (IQR) or n (%). Of the total 2 747 409 women and 2 684 730 men, BMI was not recorded in 932 720 (34.0%) women and 1 233 196 (45.9%) men, and ethnicity was not recorded in 1 278 931 (46.6%) women and 1 494 450 (55.7%) men (assumed White). \*Numbers in this row are excluded from the denominator when calculating percentages for the other rows within this section.

**Table 1: Baseline data in the derivation and validation cohorts**

	Women		Men	
	CFracture	QFracture	CFracture	QFracture
<b>Major osteoporotic fracture</b>				
All patients	0.813 (0.810–0.816)	0.813 (0.810–0.817)	0.734 (0.729–0.740)	0.736 (0.730–0.741)
Age group				
30–64 years	0.711 (0.705–0.717)	0.709 (0.703–0.715)	0.617 (0.609–0.624)	0.619 (0.611–0.627)
65–74 years	0.620 (0.613–0.627)	0.614 (0.607–0.621)	0.654 (0.641–0.666)	0.653 (0.641–0.666)
75–84 years	0.610 (0.604–0.617)	0.617 (0.610–0.623)	0.640 (0.628–0.652)	0.646 (0.635–0.658)
85–99 years	0.564 (0.555–0.574)	0.568 (0.558–0.577)	0.622 (0.601–0.643)	0.623 (0.602–0.643)
CCI				
0	0.797 (0.792–0.801)	0.795 (0.791–0.800)	0.665 (0.657–0.673)	0.665 (0.657–0.673)
1	0.801 (0.794–0.808)	0.802 (0.796–0.809)	0.725 (0.714–0.737)	0.725 (0.714–0.737)
2	0.754 (0.745–0.763)	0.753 (0.744–0.762)	0.730 (0.714–0.745)	0.730 (0.714–0.745)
≥3	0.701 (0.690–0.711)	0.711 (0.701–0.721)	0.714 (0.698–0.730)	0.714 (0.698–0.730)
<b>Hip fracture</b>				
All patients	0.914 (0.908–0.919)	0.917 (0.912–0.923)	0.886 (0.877–0.895)	0.888 (0.879–0.897)
Age group				
30–64 years	0.821 (0.805–0.837)	0.835 (0.819–0.851)	0.769 (0.751–0.787)	0.773 (0.754–0.791)
65–74 years	0.683 (0.671–0.695)	0.695 (0.683–0.707)	0.695 (0.676–0.714)	0.696 (0.677–0.715)
75–84 years	0.644 (0.636–0.653)	0.658 (0.649–0.666)	0.681 (0.666–0.695)	0.688 (0.673–0.702)
85–99 years	0.579 (0.567–0.590)	0.601 (0.589–0.612)	0.616 (0.591–0.640)	0.633 (0.608–0.657)
CCI				
0	0.922 (0.913–0.930)	0.924 (0.915–0.932)	0.848 (0.834–0.863)	0.850 (0.835–0.865)
1	0.893 (0.882–0.903)	0.897 (0.886–0.907)	0.870 (0.852–0.888)	0.872 (0.854–0.890)
2	0.833 (0.820–0.847)	0.841 (0.828–0.854)	0.800 (0.779–0.820)	0.806 (0.785–0.827)
≥3	0.767 (0.753–0.782)	0.783 (0.769–0.797)	0.774 (0.754–0.795)	0.780 (0.760–0.801)

Data are Harrell's C-statistic (95% CI). CCI=Charlson Comorbidity Index.

**Table 2: Discrimination of CFracture and QFracture in women and men in the validation cohort for major osteoporotic fracture and hip fracture**

humeral fractures identified from codes in the primary care record (using Read Codes), or Hospital Episodes Statistics discharge diagnoses or Office for National Statistics death registration (International Classification of Diseases [10th revision] codes). Codes used to define these outcomes are specified in a previous publication.<sup>11</sup>

**Prediction model**

Variables were included from the QFracture and CFracture model as outlined in the appendix (p 3). In

contrast to QFracture derivation and internal validation, which used data on BMI, alcohol, and smoking status recorded after the date of cohort entry but before any fracture outcome in prediction, we restricted predictor values to those recorded before cohort entry.

**Comorbidity**

For each patient at baseline, a Charlson Comorbidity Index (CCI) score was additionally calculated on the basis of Read Codes.<sup>14</sup> CCI category (defined as 0, 1, 2, or ≥3) was included in the competing risk model as a predictor of competing mortality risk.

**Missing data**

Missing data were managed as detailed in the appendix (p 4). Individuals with missing ethnicity were assumed to be White (as for QFracture derivation). For missing BMI, smoking status, and alcohol status, multivariate imputation by chained equations<sup>15</sup> was used to generate five imputed datasets that were combined using Rubin's rules. Morbidities and prescription medicines used for prediction were assumed to be absent if not recorded (similarly to QFracture's derivation).<sup>3</sup>

**Statistical analysis**

No formal power calculation was done because the study size was determined by the data available in CPRD, which was greater than in QFracture derivation. We implemented the published QFracture-2012 risk model (under GNU Lesser General Public Licence, version 3) and calculated QFracture-predicted 10-year risk of a major osteoporotic fracture and hip fracture, using coefficients and baseline hazard based on QFracture's original derivation cohort. Patients were randomly allocated to a fixed derivation and test dataset in a 2:1 ratio, with the split balanced in terms of age and outcome status. The derivation dataset was used to derive CFracture, a Fine-Gray model to predict the 10-year risk of a major osteoporotic fracture or hip fracture event accounting for the competing risk of non-fracture death. Separate models were estimated for men and women. Fine-Gray models calculate the subdistribution hazard ratio (ie, the instantaneous risk of a fracture event in individuals who have not yet experienced a fracture event), simultaneously accounting for the occurrence of non-fracture death.<sup>16</sup> Because we wished to explicitly compare prediction in a model accounting for competing risk versus QFracture, we included the same main effects and age interactions as in QFracture, using the same fractional polynomials for age and BMI. However, in CFracture, we also accounted for non-fracture death as a second (competing) outcome and included the CCI score as a validated predictor of mortality.<sup>17</sup> These models allow the cumulative incidence function or probability of a fracture outcome occurring over time to be directly estimated. The proportional hazards assumption was assessed by plots of Schoenfeld residuals against time

for each predictor, and by fitting and testing time-dependent terms; no evidence was found to reject this assumption.

The performance of CFracture was compared with QFracture in the independent validation dataset by examining the new model's calibration (ie, how closely the predicted and observed probabilities agree) and discrimination (ie, the ability to differentiate those who experience the outcome during the study from those who do not, expressed through Harrell's C-statistic). A C-statistic of 0·5 indicates discrimination that is no better than chance, whereas a C-statistic of 1 indicates perfect discrimination.<sup>18</sup> Calibration was evaluated by plotting the observed versus predicted risk for CFracture and QFracture. Observed risk was estimated using the Aalen-Johansen estimator, which accounts for competing mortality risk.<sup>16,17</sup> Plots were generated separately by sex, for all patients and for prespecified subgroups of age and CCI on the basis of summary statistics pooled across the imputed dataset. We also calculated a quantitative summary measure of calibration using the ratio of observed 10-year risk to mean predicted risk, which assesses how close the predicted risk is to the overall observed outcome proportion. A ratio of 1 indicates perfect calibration; a ratio of less than 1 indicates that, on average, the model overpredicts; and a ratio of more than 1 indicates that, on average, the model underpredicts. Discrimination and calibration were additionally measured using a complete case analysis in the validation cohort.

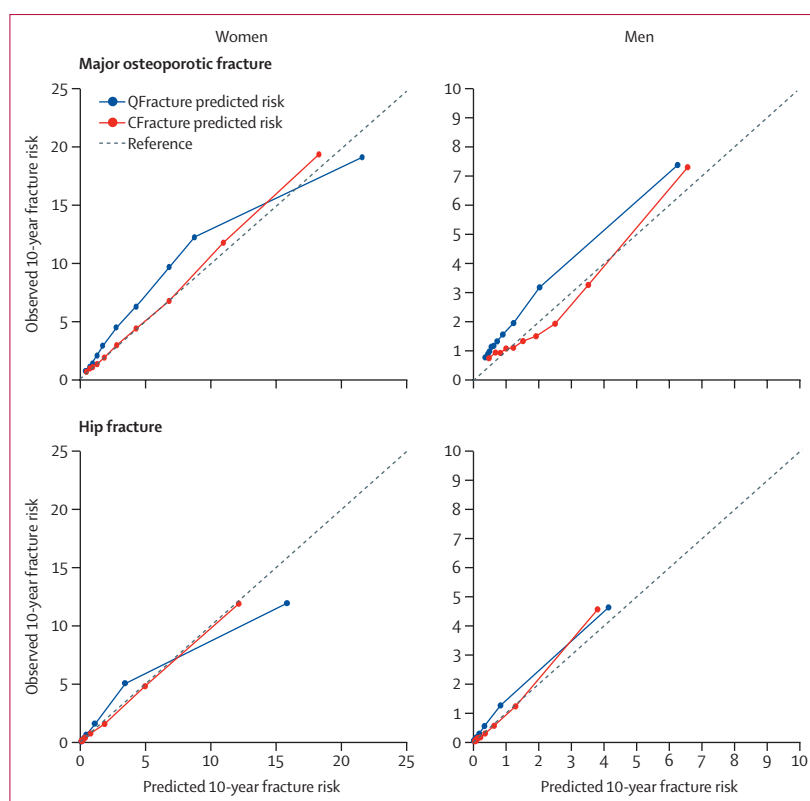
We also examined changes in patients reclassified either side of several potential thresholds of 10-year fracture risk—namely 5%, 10%, and 20%. To visualise reclassification, we generated Sankey diagrams and scatterplots of risk predicted by CFracture versus QFracture. We examined the characteristics of reclassified patients and estimated the number needed to treat (NNT) to prevent one new major osteoporotic fracture or hip fracture, assuming that all people recommended for treatment take an osteoporosis treatment, using a hypothetical relative risk reduction of 20% and 40% for new fracture events. The NNT was calculated as:

$$\frac{\text{Number of fractures} \times \text{hypothetical proportional risk reduction}}{\text{Number of patients recommended for treatment}}$$

Analyses were conducted using R version 4.2.0. Specific R packages and versions included mice 3.14.0, survival 3.3.1, prodlim 2019.11.13, rms 6.3.0, cmprsk 2.2.10, riskRegression 2022.3.22, pec 2021.10.11, splines 4.2.0, and geePack 1.3.2.

#### Role of the funding source

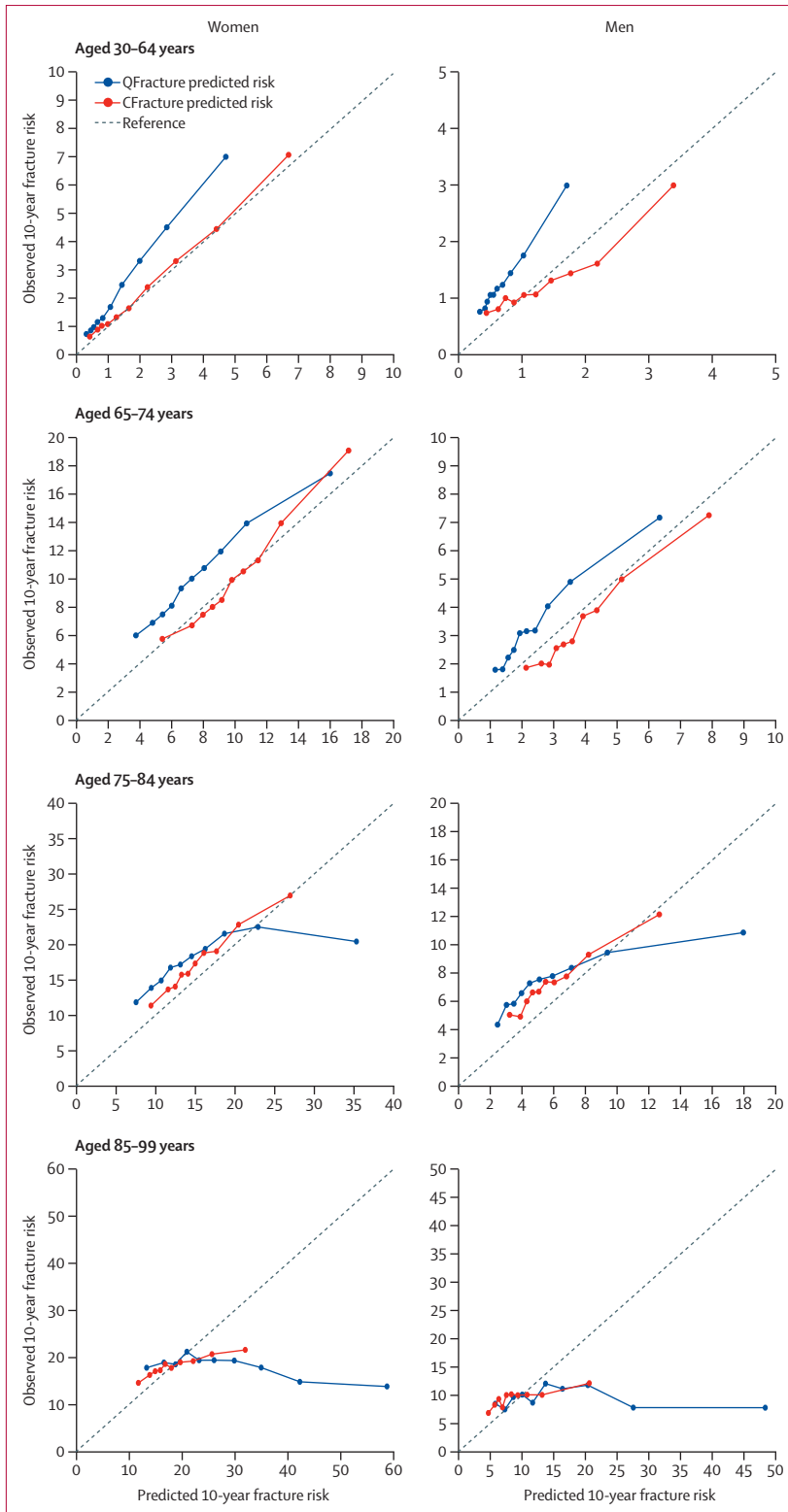
The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.



**Figure 1:** Overall calibration of CFracture and QFracture for major osteoporotic fracture and hip fracture. Observed risk is based on the Aalen-Johansen estimator, which accounts for competing mortality risk. Ideal calibration lies on the reference line, below the line is overprediction, and above the line is underprediction.

## Results

The derivation cohort included 1831606 women and 1789820 men, and the validation cohort included 915803 women and 894910 men. Baseline characteristics were similarly distributed in the derivation and validation cohorts (table 1). The number of women with missing data ranged from 441510 (16·1%) for smoking status to 1278931 (46·6%) for ethnicity, and the number of men with missing data ranged from 705402 (26·3%) for smoking status to 1494450 (55·7%) for ethnicity. At 10 years of follow-up in the derivation cohort, 63783 major osteoporotic fractures were recorded in women over 10419774 person-years of follow-up (incidence 6·12 [95% CI 6·07–6·17] per 1000 person-years), compared with 22863 cases in men over 10123610 person-years of follow-up (incidence 2·26 [2·23–2·29] per 1000 person-years); and 26782 hip fractures were recorded in women over 11446139 person-years of follow-up (incidence 2·34 [2·31–2·37] per 1000 person-years), compared with 9976 cases in men over 10984254 person-years of follow-up (incidence 0·91 [0·89–0·93] per 1000 person-years). Major osteoporotic fracture and hip fracture incidence rose progressively with age (appendix p 5). The sex-specific Fine-Gray models for major osteoporotic fracture and hip fracture are shown in the appendix (pp 6–13).



**Figure 2: Calibration of CFracture and QFracture for major osteoporotic fracture by age group**  
 Observed risk is based on the Aalen-Johansen estimator, which accounts for competing mortality risk. Ideal calibration lies on the reference line, below the line is overprediction, and above the line is underprediction.

In the validation cohort, discrimination of CFracture was very similar to QFracture discrimination in the overall population cohort. For major osteoporotic fracture, CFracture was excellent in women and good in men (in women: C-statistic 0·813 [95% CI 0·810–0·816] for CFracture vs 0·813 [0·810–0·817] for QFracture; in men: 0·734 [0·729–0·740] for CFracture vs 0·736 [0·730–0·741] for QFracture). For hip fracture, CFracture discrimination was excellent in both women and men (in women: C-statistic 0·914 [0·908–0·919] for CFracture vs 0·917 [0·912–0·923] for QFracture; in men: 0·886 [0·877–0·895] for CFracture vs 0·888 [0·879–0·897] for QFracture; table 2). Discrimination of CFracture varied by age group and CCI stratum in women, and was best in the younger women (30–64 years age group) and in the least comorbid groups (CCI 0–2) and worst in the oldest women (age groups >64 years) and in the most comorbid groups (CCI ≥3). For hip fracture, discrimination in men generally followed a similar pattern as in women, being best in younger men (30–64 years age group) and in the least comorbid groups (CCI 0–2), and worst in the oldest men (85–99 years age group) and in the most comorbid groups (CCI ≥3). However, for major osteoporotic fracture, discrimination of CFracture and QFracture in men was poor for all age groups, with some improvement in higher-comorbidity groups (CCI ≥1).

Calibration of CFracture for major osteoporotic fracture was better than QFracture in the whole population cohort, with some underprediction in women at higher levels of predicted risk. Calibration of CFracture for major osteoporotic fracture in men was similar, but additionally showed some overprediction in the middle range of predicted risk (figure 1). The ratio of observed 10-year risk to mean predicted risk in both men and women was closer to 1 with CFracture than with QFracture for prediction of major osteoporotic fracture (appendix p 14). Calibration of CFracture by sex and age group is shown in figure 2, and calibration by CCI stratification is shown in figure 3.

Calibration of CFracture for hip fracture was better than that of QFracture in the whole population cohort. Calibration was good in women across all levels of predicted risk, and good in men except in the highest decile of predicted risk, where there was underprediction (figure 1). The ratio of observed 10-year risk to mean predicted risk in women was closer to 1 with CFracture than with QFracture for prediction of hip fracture, but was similar in men (appendix p 14). Calibration was good in women aged 30–64 and 75–84 years, with some overprediction in women aged 65–74 years (appendix p 15). In men, calibration was reasonable in those aged 30–64, 65–74, and 75–84 years, with some underprediction at the highest level of predicted risk. In both men and women aged 85–99 years, calibration was poor, with overprediction at most levels of predicted risk, although it was considerably better than QFracture in this age group as well as all others. Stratified by CCI, calibration in women was good except in the highest

decile of predicted risk, where there was some overprediction in those with a CCI of 1, 2, or 3 or higher (appendix p 16). Calibration in men stratified by CCI was good apart from underprediction in the highest decile of predicted risk for a CCI of 0 or 1, underprediction in the middle of the predicted risk range for a CCI of 2, and some overprediction in the highest decile of predicted risk for a CCI of 3 or higher.

The proportion of people reclassified by CFrature compared with QFrature above and below each risk threshold examined is shown in the appendix (pp 17–20). For major osteoporotic fracture in women, net reclassification up with QFrature occurred for those with a predicted risk below 15%, and net reclassification down occurred for those at or above 15%. For major osteoporotic fracture in men, net reclassification down occurred for those with a QFrature predicted risk at or above 10% and little movement for lower risk. For hip fracture, reclassification down was higher than reclassification up in women with a QFrature predicted risk at or above 10%, and across all levels of QFrature predicted risk in men. Scatterplots of CFrature versus QFrature predicted risk were in keeping with these differences (appendix p 21).

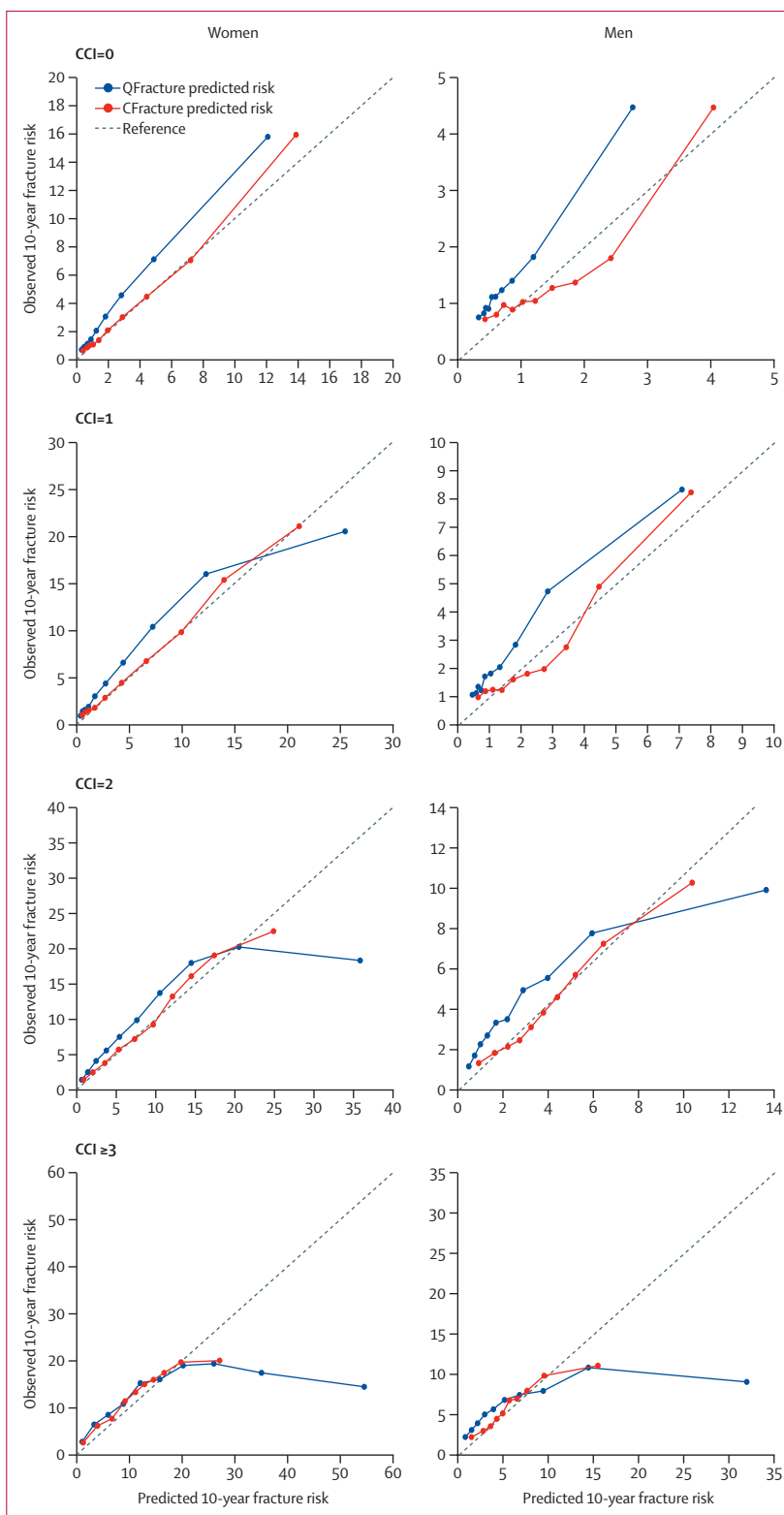
The number of patients classified at or above potential thresholds, number of fractures, and estimated NNT to prevent one fracture are shown in table 3. At a 20% risk threshold, CFrature had a lower estimated NNT than QFrature (eg, major osteoporotic fracture in women: NNT 30·6 [95% CI 28·7–32·7] vs 38·6 [36·5–40·9] with 20% treatment effectiveness). At a 10% risk threshold, CFrature had a similar NNT in women and a lower NNT in men than QFrature. At a 5% threshold, CFrature had a higher estimated NNT for both major osteoporotic fracture and hip fracture (except for hip fracture in men, where CFrature and QFrature produced similar NNT estimates).

Compared with women classified at or above a 10% risk threshold for major osteoporotic fracture by QFrature, women reclassified at or above a 10% risk threshold for major osteoporotic fracture by CFrature were younger and had a higher prevalence of current smoking and previous fracture, but a lower prevalence of dementia, cancer, cardiovascular disease, renal disease, and type 2 diabetes (appendix p 22). This pattern was reproduced in men, apart from current smoking status, which was similar in both groups. The pattern was similar for hip fracture but with smaller differences in these covariates (appendix p 23).

Discrimination and calibration of both QFrature and CFrature in the overall population cohort with complete case analysis was similar to the main results (appendix pp 24–25).

## Discussion

Discrimination of CFrature in the whole population validation cohort ranged from good to excellent for



**Figure 3: Calibration of CFrature and QFrature for MOF by comorbidity group**

Observed risk is based on the Aalen-Johansen estimator, which accounts for competing mortality risk. Ideal calibration lies on the reference line, below the line is overprediction, and above the line is underprediction. CCI=Charlson Comorbidity Index.



	Number recommended for treatment	Events	NNT (95% CI)	
			20% relative risk reduction	40% relative risk reduction
<b>Women, major osteoporotic fracture</b>				
20% risk threshold				
QFracture	44 307	5746	38.6 (36.5–40.9)	19.3 (18.5–20.1)
CFracture	26 293	4298	30.6 (28.7–32.7)	15.3 (14.6–16.0)
10% risk threshold				
QFracture	119 993	15 209	39.4 (38.1–40.9)	19.7 (19.3–20.2)
CFracture	156 298	18 521	42.2 (40.9–43.6)	21.1 (20.6–21.6)
5% risk threshold				
QFracture	217 761	22 036	49.4 (48.0–50.9)	24.7 (24.2–25.2)
CFracture	289 563	25 037	57.8 (56.3–59.5)	28.9 (28.4–29.5)
<b>Women, hip fracture</b>				
20% risk threshold				
QFracture	29 202	2534	57.6 (53.0–63.1)	28.8 (27.2–30.7)
CFracture	4576	448	51.1 (42.4–64.2)	25.5 (22.3–29.8)
10% risk threshold				
QFracture	64 713	5763	56.1 (53.1–59.6)	28.1 (27.0–29.2)
CFracture	61 694	5696	54.2 (51.2–57.5)	27.1 (26.0–28.2)
5% risk threshold				
QFracture	109 784	8904	61.6 (58.9–64.6)	30.8 (29.9–31.9)
CFracture	134 200	10 066	66.7 (63.9–69.7)	33.3 (32.3–34.4)
<b>Men, major osteoporotic fracture</b>				
20% risk threshold				
QFracture	5869	347	84.6 (68.5–110.4)	42.3 (36.3–50.6)
CFracture	1220	122	50.0 (35.9–82.3)	25.0 (19.6–34.5)
10% risk threshold				
QFracture	18 625	1227	75.9 (67.5–86.7)	37.9 (34.9–41.6)
CFracture	11 102	890	62.4 (54.4–73.0)	31.2 (28.3–34.7)
5% risk threshold				
QFracture	47 889	2865	83.6 (77.3–91.0)	41.8 (39.5–44.3)
CFracture	67 136	3745	89.6 (83.7–96.5)	44.8 (42.7–47.2)
<b>Men, hip fracture</b>				
20% risk threshold				
QFracture	5030	230	109.3 (84.9–153.5)	54.7 (45.5–68.6)
CFracture	414	29	71.4 (39.5–372.0)	35.7 (22.8–82.5)
10% risk threshold				
QFracture	13 111	631	103.9 (88.5–125.7)	51.9 (46.3–59.2)
CFracture	4695	274	85.7 (67.8–116.3)	42.8 (36.1–52.6)
5% risk threshold				
QFracture	29 992	1394	107.6 (96.3–121.8)	53.8 (49.7–58.6)
CFracture	26 872	1287	104.4 (93.1–118.8)	52.2 (48.1–57.1)

NNT values are calculated assuming either a 20% or 40% relative risk reduction with a hypothetical osteoporosis treatment taken by all people recommended for treatment. NNT=number needed to treat with osteoporosis treatment to prevent one fracture.

**Table 3: Number of people classified at or above a percentage threshold by CFracture and QFracture (and therefore recommended for treatment), observed events, and estimated number needed to treat with a hypothetical osteoporosis treatment to prevent an incident fracture event**

major osteoporotic fracture and was excellent for hip fracture, similarly to QFracture discrimination. Worse discrimination occurred with increasing age except for

major osteoporotic fracture in men, for whom no obvious relationship with age was noted. Stratified by CCI, no clear pattern with increasing CCI for discrimination of major osteoporotic fracture was noted, whereas discrimination for hip fracture was similar to the whole population cohort for a CCI of 0 and declined with increasing comorbidity (although it was good or excellent in all strata). CFracture was better calibrated than QFracture in the whole population cohort and in every stratum, although it was better calibrated in women than in men, and for hip fracture than for major osteoporotic fracture. CFracture was poorly calibrated in both women and men aged 85–99 years (although less so than QFracture).

The first version of QFracture was externally validated in The Health Improvement Network (a UK database), and the updated version of QFracture was validated in CPRD-GOLD. Both validations showed excellent discrimination and calibration in the whole population cohort.<sup>19,20</sup> However, these evaluated QFracture in the context of incomplete fracture ascertainment because fractures were only identified in primary care records (with or without linked mortality data) and did not include hospitalisations, as reflected in their lower fracture incidence, potentially leading to underprediction of fracture risk. QFracture has not been validated across age groups or in those with comorbidity, in whom we show it has worse calibration than a model accounting for competing mortality risk, consistent with systematic overprediction by QFracture.

Other than FRAX, few tools account for competing mortality risk in the context of fracture prediction. Two studies were conducted in older people (aged ≥60 years): one to quantify the effects of the predictors on the transition risks to fracture and mortality over 5 years, and the other to quantify residual lifetime fracture risk.<sup>21,22</sup> One study estimated fracture risk in patients with type 2 diabetes, showing that failing to account for competing risk mortality overestimates fracture risk over a 5-year time horizon.<sup>23</sup> Another competing risk model used to predict fracture risk showed good discrimination and calibration, but was only derived and validated for use in post-menopausal women.<sup>24</sup>

Despite similar discrimination to QFracture, CFracture reclassified different patients around thresholds of predicted risk. Compared with QFracture, CFracture classified more women below a potential 10% risk threshold when predicting major osteoporotic fracture, similar numbers of women when predicting hip fracture, and fewer men for both major osteoporotic fracture and hip fracture. At a potential 20% risk threshold, CFracture also tended to have a lower estimated NNT for a 20% and 40% hypothetical treatment effectiveness. At lower risk thresholds (ie, 5% and 10%), differences tended to be smaller. CFracture recommended treatment in fewer patients with a history of comorbidities associated with death (as seen, for example, with the lower prevalence of

dementia and cancer among patients recommended to have treatment), who will have a greater incidence of competing mortality risk. We show that, irrespectively of whether such risk scores have similar overall discrimination, different patients are classified by each tool around thresholds (similar to what is observed in cardiovascular risk prediction),<sup>9</sup> which could be important if such characteristics are associated with the likelihood of treatment response or adherence to therapy.

Our study was robust and used data from a large, representative population. We included linked data for better fracture outcome ascertainment, which partly accounts for the differences in hip fracture and major osteoporotic fracture incidence seen in previous QFracture derivation and validation studies published in peer-reviewed journals.<sup>3</sup> However, our study has limitations. We only allowed data collected before cohort entry to be used in prediction (because using future data in prediction models can lead to bias), which meant our study has greater amounts of missing data than those in the QFracture derivation. We have provided all codes used to define variables and all model estimates (unlike FRAX, for example, for which the algorithm has never been published), facilitating transparency and reproducibility; however, we could not compare the Read codes we used to define variables with QFracture because QFracture code lists were not publicly available at the time. Although we explicitly accounted for censoring due to death, our model assumes that people who deregister from a general practice have the same fracture risk as those who do not. However, deregistration from a general practice might be informative—eg, if some older people with many risk factors for fracture move into residential care. We are unable to identify subclinical vertebral fractures, which is true for any other fracture model using electronic health records or administrative health data, including QFracture. As has been done with QFracture and elsewhere, we used multiple imputation for these missing data, which relies on the assumption that all data are missing at random.<sup>3</sup> We had a higher proportion of missing data than in QFracture derivation because we avoided including forward-looking values in prediction. Despite this, sensitivity analysis using a complete case analysis showed results that were similar to those from the main analysis. For other variables, we assumed that the condition, or family history, was not present if it was not recorded. We also used a later index date for cohort entry than QFracture, because we wished to better account for improved data recording in electronic health records and because using increasingly historical data to derive clinical prediction tools might result in bias.<sup>25</sup> CFracture was derived and validated in the same dataset (internally validated), whereas QFracture is externally validated, having been derived from a different dataset. Although the performance of QFracture in the derivation dataset and in CPRD-GOLD have been shown to be comparable, further external validation of CFracture

is required.<sup>26</sup> Fracture incidence increases with age and QFracture is recommended to be used in people aged 30–84 years. We found that QFracture and CFracture performed poorly in patients aged 85 years and older, in keeping with this guidance. This suggests that other risk factors that are not currently accounted for in this population might exist. The thresholds we used to estimate the NNT were arbitrary. A 10% threshold is the threshold at which intravenous bisphosphonates are considered cost-effective by NICE.<sup>27</sup> The UK National Osteoporosis Guideline Group uses age-dependent thresholds to guide bone mineral density measurement until the age of 70 years, at which point a fixed lower (11.1%) and upper (20.3%) threshold is recommended.<sup>5</sup> We also examined reclassification around a 5% threshold to explore impact if future guidelines were to recommend larger numbers of people for treatment (eg, as drug costs fall, the cost-effectiveness threshold also falls). CPRD-GOLD also does not include all general practices in England, and small numbers of people might not be registered with a general practice. Data collected within electronic health records might also be influenced by patients in poorer health, who are more likely to interact with health professionals.

We reported prediction over a 10-year period to show how accounting for competing mortality risk and including a predictor of non-fracture death improves performance compared with a nationally recommended model for 10-year fracture prediction. Other tools also predict fracture risk using a 10-year period, including FRAX and the Garvan Fracture Risk Calculator. This time period is likely to cover the duration of osteoporosis treatment and potential benefits that might continue if treatment is discontinued. Although shorter time periods have been reported to have limited benefit in aiding risk categorisation, they have been suggested to be considered in older adults.<sup>28,29</sup> NICE recommends using either QFracture or FRAX to assess fracture risk, but recognises that they are designed differently and are not interchangeable. Since its release, FRAX models have been validated in many countries, but independent external validation is challenging at scale because the FRAX equation is not published.<sup>30</sup> FRAX accounts for competing risk and can be used with or without bone mineral density. However, FRAX has been shown to overpredict fracture risk when an incomplete method of fracture ascertainment was used<sup>31</sup> and to underestimate major osteoporotic fracture risk in studies with more complete fracture ascertainment.<sup>32</sup> FRAX has also been criticised for only using binary clinical risk factors that do not account for exposure response.<sup>33</sup> This makes robust head-to-head comparisons between CFracture and FRAX difficult (although they should still be performed). The Garvan calculator has also been used to individualise the risk of fragility fractures over 10 years.<sup>6</sup> Garvan includes risk factors such as history of previous fracture, history of fall during the past 12 months, age,

and bone mineral density, but does not account for competing risk. Omnibus models predicting for very wide age ranges will always have good discrimination because age is such a strong predictor, but good discrimination does not mean accurate prediction, as shown by differences in calibration. Therefore, competing risks should be routinely considered if prediction models are aimed at older populations with comorbidity.

In conclusion, QFracture underpredicted fracture risk in young and healthy individuals, and overpredicted fracture risk in older individuals and those with comorbidity. A tool such as CFracture that accounts for competing risk should therefore be considered for clinical use.

#### Contributors

BG, DRM, and PTD conceived and designed the study and obtained the funding. All authors contributed to study design and interpretation. SJL, MM, CE, and DRM led the data management. SJL led the analysis, supported by BG, DRM, and PTD. SJL and DRM drafted the paper, which all authors revised. SJL, BG, and DRM verified the underlying data. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

#### Declaration of interests

PTD reports a grant from AbbVie, outside the submitted work, and is a member of the NHS Scottish Medicines Consortium. DRM reports grants from the Chief Scientist Office, Health Data Research UK, and National Institute for Health and Care Research (NIHR), outside the submitted work. DRM was a member of the European Medicines Agency Pharmacovigilance Risk Assessment Committee. All other authors declare no competing interests.

#### Data sharing

The data controller is the Clinical Practice Research Datalink and, under the data licence granted, we are not allowed to share data. Researchers can apply to the Clinical Practice Research Datalink directly for access to the raw data (for further details see <https://cprd.com/>).

#### Acknowledgments

This study was funded by the NIHR Health and Social Care Delivery Research Programme (project reference 15/12/22). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care. The authors had full and sole access to the data, and the funder had no role in the conduct of the research or the decision to publish. DRM is supported by a Wellcome Trust Clinical Research Development Fellowship (grant 214588/Z/18/Z). CE is funded by Legal and General (philanthropic support for the Advanced Care Research Centre). SJL was part funded by Health Data Research UK.

#### References

- Shen Y, Huang X, Wu J, et al. The global burden of osteoporosis, low bone mass, and its related fracture in 204 countries and territories, 1990-2019. *Front Endocrinol (Lausanne)* 2022; 13: 882241.
- National institute for Health and Care Excellence. Osteoporosis—assessing the risk of fragility fractures [CG146]. <https://www.nice.org.uk/guidance/cg146> (accessed June 14, 2022).
- Hippisley-Cox J, Coupland C. Derivation and validation of updated QFracture algorithm to predict risk of osteoporotic fracture in primary care in the United Kingdom: prospective open cohort study. *BMJ* 2012; 344: e3427.
- McCloskey EV, Johansson H, Oden A, Kanis JA. From relative risk to absolute fracture risk calculation: the FRAX algorithm. *Curr Osteoporos Rep* 2009; 7: 77–83.
- Gregson CL, Armstrong DJ, Bowden J, et al. UK clinical guideline for the prevention and treatment of osteoporosis. *Arch Osteoporos* 2022; 17: 58.
- Nguyen ND, Frost SA, Center JR, Eisman JA, Nguyen TV. Development of a nomogram for individualizing hip fracture risk in men and women. *Osteoporos Int* 2007; 18: 1109–17.
- Wolbers M, Koller MT, Witteman JCM, Steyerberg EW. Prognostic models with competing risks: methods and application to coronary risk prediction. *Epidemiology* 2009; 20: 555–61.
- Livingstone S, Morales DR, Donnan PT, et al. Effect of competing mortality risks on predictive performance of the QRISK3 cardiovascular risk prediction tool in older people and those with comorbidity: external validation population cohort study. *Lancet Healthy Longev* 2021; 2: e352–61.
- Livingstone SJ, Guthrie B, Donnan PT, Thompson A, Morales DR. Predictive performance of a competing risk cardiovascular prediction tool CRISK compared to QRISK3 in older people and those with comorbidity: population cohort study. *BMC Med* 2022; 20: 152.
- Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. *Circulation* 2016; 133: 601–09.
- Livingstone SJ, Morales DR, McMinn M, Eke C, Donnan P, Guthrie B. Effect of competing mortality risks on predictive performance of the QFracture risk prediction tool for major osteoporotic fracture and hip fracture: external validation cohort study in a UK primary care population. *BMJ Med* 2022; 1: e000316.
- Herrett E, Gallagher AM, Bhaskaran K, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015; 44: 827–36.
- Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the General Practice Research Database: a systematic review. *Br J Clin Pharmacol* 2010; 69: 4–14.
- Khan NF, Perera R, Harper S, Rose PW. Adaptation and validation of the Charlson Index for Read/OXMIS coded databases. *BMC Fam Pract* 2010; 11: 1.
- van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011; 45: 67.
- Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med* 2007; 26: 2389–430.
- van Geloven N, Giardiello D, Bonneville EF, et al. Validation of prediction models in the presence of competing risks: a guide through modern methods. *BMJ* 2022; 377: e069249.
- Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010; 21: 128–38.
- Collins GS, Mallett S, Altman DG. Predicting risk of osteoporotic and hip fracture in the United Kingdom: prospective independent and external validation of QFractureScores. *BMJ* 2011; 342: d3651.
- Dagan N, Cohen-Stavi C, Leventer-Roberts M, Balicer RD. External validation and comparison of three prediction tools for risk of osteoporotic fractures using data from population based electronic health records: retrospective cohort study. *BMJ* 2017; 356: i6755.
- Nguyen ND, Ahlborg HG, Center JR, Eisman JA, Nguyen TV. Residual lifetime risk of fractures in women and men. *J Bone Miner Res* 2007; 22: 781–88.
- Tran T, Bliuc D, Pham HM, et al. A risk assessment tool for predicting fragility fractures and mortality in the elderly. *J Bone Miner Res* 2020; 35: 1923–34.
- Tebé C, Martínez-Laguna D, Moreno V, et al. Differential mortality and the excess rates of hip fracture associated with type 2 diabetes: accounting for competing risks in fracture prediction matters. *J Bone Miner Res* 2018; 33: 1417–21.
- Baleanu F, Moreau M, Charles A, et al. Fragility fractures in postmenopausal women: development of 5-year prediction models using the FRISBEE study. *J Clin Endocrinol Metab* 2022; 107: e2438–48.
- Pylypchuk R, Wells S, Kerr A, et al. Cardiovascular disease risk prediction equations in 400000 primary care patients in New Zealand: a derivation and validation study. *Lancet* 2018; 391: 1897–907.
- Hippisley-Cox J, Coupland C, Brindle P. The performance of seven QFracture risk scores in an independent external sample of patients from general practice: a validation study. *BMJ Open* 2014; 4: e005809.
- National Institute for Health and Care Excellence. Bisphosphonates for treating osteoporosis: Technology appraisal guidance [TA464]. <https://www.nice.org.uk/guidance/ta464/chapter/3-Committee-discussion> (accessed June 16, 2022).

- 
- 28 Kanis JA, Johansson H, Harvey NC, et al. The use of 2-, 5-, and 10-year probabilities to characterize fracture risk after a recent sentinel fracture. *Osteoporos Int* 2021; **32**: 47–54.
- 29 McCloskey EV, Borgstrom F, Cooper C, et al. Short time horizons for fracture prediction tools: time for a rethink. *Osteoporos Int* 2021; **32**: 1019–25.
- 30 Marques A, Ferreira RJ, Santos E, Loza E, Carmona L, da Silva JA. The accuracy of osteoporotic fracture risk prediction tools: a systematic review and meta-analysis. *Ann Rheum Dis* 2015; **74**: 1958–67.
- 31 Hippisley-Cox J, Coupland C. Validation of QFracture compared with FRAX: analysis prepared for NICE 2011. Nottingham: University of Nottingham, 2011. <https://www.qresearch.org/media/1033/a-validation-of-qfracture-vs-frax-for-nice-2011-13.pdf> (accessed Aug 13, 2022).
- 32 Baleanu F, Iconaru L, Charles A, et al. Independent external validation of FRAX and Garvan Fracture Risk Calculators: a sub-study of the FRISBEE Cohort. *JBMR Plus* 2021; **5**: e10532.
- 33 Vandenput L, Johansson H, McCloskey EV, et al. Update of the fracture risk prediction tool FRAX: a systematic review of potential cohorts and analysis plan. *Osteoporos Int* 2022; **33**: 2103–36.