

Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores

Julia Hippisley-Cox, professor of clinical epidemiology and general practice, Carol Coupland, associate professor in medical statistics

Division of Primary Care, Tower Building, University Park, Nottingham NG2 7RD

Correspondence to: J Hippisley-Cox
Julia.hippisley-cox@nltworld.com

Cite this as: *BMJ* 2009;339:b4229
doi:10.1136/bmj.b4229

ABSTRACT

Objective To develop and validate two new fracture risk algorithms (QFractureScores) for estimating the individual risk of osteoporotic fracture or hip fracture over 10 years.

Design Prospective open cohort study with routinely collected data from 357 general practices to develop the scores and from 178 practices to validate the scores.

Setting General practices in England and Wales.

Participants 1 183 663 women and 1 174 232 men aged 30-85 in the derivation cohort, who contributed 7 898 208 and 8 049 306 person years of observation, respectively. There were 24 350 incident diagnoses of osteoporotic fracture in women and 7934 in men, and 9302 incident diagnoses of hip fracture in women and 5424 in men.

Main outcome measures First (incident) diagnosis of osteoporotic fracture (vertebral, distal radius, or hip) and incident hip fracture recorded in general practice records.

Results Use of hormone replacement therapy (HRT), age, body mass index (BMI), smoking status, recorded alcohol use, parental history of osteoporosis, rheumatoid arthritis, cardiovascular disease, type 2 diabetes, asthma, tricyclic antidepressants, corticosteroids, history of falls, menopausal symptoms, chronic liver disease, gastrointestinal malabsorption, and other endocrine disorders were significantly and independently associated with risk of osteoporotic fracture in women. Some variables were significantly associated with risk of osteoporotic fracture but not with risk of hip fracture. The predictors for men for osteoporotic and hip fracture were age, BMI, smoking status, recorded alcohol use, rheumatoid arthritis, cardiovascular disease, type 2 diabetes, asthma, tricyclic antidepressants, corticosteroids, history of falls, and liver disease. The hip fracture algorithm had the best performance among men and women. It explained 63.94% of the variation in women and 63.19% of the variation in men. The D statistic values for discrimination were highest for hip fracture in women (2.73) and men (2.68) and were over twice the magnitude of the corresponding values for osteoporotic fracture. The ROC statistics for hip fracture were also high: 0.89 in women and 0.86 for men versus 0.79 and 0.69, respectively, for the osteoporotic fracture outcome. The algorithms were well calibrated with predicted risks

closely matching observed risks. The QFractureScore for hip fracture also had good performance for discrimination and calibration compared with the FRAX (fracture risk assessment) algorithm.

Conclusions These new algorithms can predict risk of fracture in primary care populations in the UK without laboratory measurements and are therefore suitable for use in both clinical settings and for self assessment (www.qfracture.org). QFractureScores could be used to identify patients at high risk of fracture who might benefit from interventions to reduce their risk.

INTRODUCTION

Osteoporotic fractures are a major and increasing cause of morbidity in the population and a considerable burden to health services. Hip fractures, in particular, result in considerable pain, loss of function, and admission to hospital, making prevention a high priority for patients and physicians and for public health. Various therapeutic and lifestyle interventions might reduce the risk of osteoporosis and hence an individual's risk of fracture.¹ The challenge now is to improve methods for accurate identification of individuals at high risk who might benefit from a therapeutic or preventive intervention. While there is no universally accepted policy for screening for patients at risk of osteoporotic fracture, some guidelines,²⁻⁵ but not all,⁶ recommend a targeted approach to the prevention of osteoporosis based on the 10 year absolute risk of major osteoporotic fracture. Risk prediction utilities are therefore required to accurately estimate individual risk as well as enable a systematic targeted population based screening approach.

Traditional approaches based on measurement of bone mineral density alone are unsuitable for population screening because of cost and low sensitivity.⁷ Most fractures occur in women with normal bone mineral density,⁸ and the evidence suggests that risk prediction algorithms that do not include bone mineral density are almost as good as those that do.⁹ We need less expensive and more practical methods of identifying those at high risk; these should ideally be based on models developed from contemporaneous data in diverse populations representative of the clinical

setting in which these models will subsequently be applied. Risk prediction utilities tend to perform best in the settings in which they have been derived both in terms of discrimination (that is, ability to separate out those who will and will not develop a fracture) and also calibration (how closely the overall predicted risk matches the observed risk).¹⁰

There are several established risk factors for osteoporotic fracture that could be used to derive a risk prediction algorithm for use within primary care. Many of these risk factors are reliably recorded within primary care clinical computer systems and hence such data can be used to derive robust utilities that can then be applied in primary care. The incidence of fracture and the prevalence of associated risk factors will change over time, and the methods to derive the risk prediction algorithms need to be dynamic so that they can be remodelled over time. UK datasets derived from family practices have the advantage of having large and broadly representative populations with historical data tracking back well over a decade in most practices, and they are continually updated.

We developed and validated two new fracture clinical risk scores (QFractureScores) derived from a large and representative primary care population from a validated clinical research database (www.qresearch.org). We analysed more than two million patients to address some of the research questions regarding risk factors for osteoporotic fracture in men and women highlighted within the recent NICE guidance,⁶ National Osteoporosis Guideline,⁵ and the World Health Organization.¹¹ We incorporated traditional variables already included in the FRAX (fracture risk assessment) algorithm⁷ and added additional variables that affect risk of fracture, such as history of falls, type 2 diabetes, cardiovascular disease, asthma, use of hormone replacement therapy (HRT), and use of tricyclic antidepressants.^{5,12-14} We also extended the age range to include younger patients. Lastly, we incorporated a more detailed categorisation of alcohol and smoking status. Our new algorithm is based on variables that are readily available in patients' electronic primary healthcare records¹⁵ or that the patients themselves would probably know, without the need for laboratory tests or clinical measurements. This approach is designed to enable the algorithms to be readily and cost effectively implemented in routine clinical practice or used by individual patients.

METHODS

Study design and data source

We conducted a prospective cohort study in a large primary care population of patients from version 20 of the QResearch database. This is a large validated primary care electronic database containing the health records of over 11 million patients registered from 574 general practices that use the Egton Medical Information System (EMIS) computer system. Practices and patients contained on the database are nationally representative for England and Wales and similar to those on other large national primary care databases that use other clinical software systems.¹⁶

Practice selection

We included all QResearch practices in England and Wales once they had been using their current EMIS system for at least a year to ensure completeness of recording of morbidity and prescribing data. We randomly allocated two thirds of practices to the derivation dataset and the remaining third to the validation dataset; the simple random sampling utility in Stata was used to assign practices to the derivation or validation cohort.

Cohort selection

We identified an open cohort of patients aged 30-85 at the study entry date, drawn from patients registered with eligible practices during the 15 years between 1 January 1993 and 30 June 2008. We used an open cohort design, rather than a closed cohort design, as this allows patients to enter the population throughout the whole study period rather than require registration on a fixed date, reflecting the realities of routine clinical practice. We excluded patients with a previous recorded fracture (hip, distal radius, or vertebral), temporary residents, patients with interrupted periods of registration with the practice, and those who did not have a valid Townsend deprivation score related to the postcode (about 4% of the population).

For each patient, we determined an entry date to the cohort, which was the latest of the date of their 30th birthday, date of registration with the practice, date on which the practice computer system was installed plus one year, and the beginning of the study period (1 January 1993). We included patients in the analysis only once they had a minimum of one year's complete data in their medical record.¹⁷ For each patient we also determined an exit date, which was the earliest of date of recorded fracture, date of death, date of deregistration with the practice, date of last upload of computerised data, or the study end date (30 June 2008).

Primary outcomes

Our two primary outcomes were the first (incident) diagnosis of an osteoporotic fracture (hip, vertebral, or distal radius) as recorded on the general practice computer records and incident diagnosis of hip fracture.

Fracture risk factors

We examined the following explanatory variables in our analysis, all of which are known or thought to affect fracture risk and are also likely to be recorded within the patients' electronic records as part of routine clinical practice:

- Age at study entry (in single years)
- Body mass index (BMI) (continuous)¹⁸
- Smoking status (non-smoker, former smoker, light smoker (<10 cigarettes/day), moderate smoker (10-19 cigarettes/day), heavy smoker (≥ 20 cigarettes/day)¹⁹⁻²¹
- Townsend deprivation score (with 2001 census data, evaluated at output area, as a continuous variable)

- Recorded parental history of osteoporosis or hip fracture in a first degree relative (binary variable yes/no)²²
- Diagnosis of cardiovascular disease at baseline (binary variable yes/no)¹²
- Recorded use of alcohol (none, trivial <1 unit/day, light 1-2 units/day, medium 3-6 units/day, heavy 7-9 units/day, very heavy >9 units/day)²³
- Diagnosis of rheumatoid arthritis at baseline (binary variable yes/no)²⁴
- Diagnosis of type 2 diabetes at baseline (binary variable yes/no)²⁵
- Diagnosis of asthma at baseline (binary variable yes/no)
- History of falls before baseline (binary variable yes/no)
- Diagnosis of chronic liver disease at baseline (binary variable yes/no)
- Diagnosis of gastrointestinal conditions likely to result in malabsorption (such as Crohn's disease, ulcerative colitis, coeliac disease, steatorrhoea, blind loop syndrome) at baseline (binary variable yes/no)⁷
- Diagnosis of other endocrine conditions (thyrotoxicosis, primary or secondary hyperparathyroidism, Cushing's syndrome) at baseline (binary variable yes/no)
- At least two prescriptions for systemic corticosteroids in the six months before baseline (binary variable yes/no)²⁶
- At least two prescriptions for tricyclic antidepressants in the six months before baseline (binary variable yes/no)¹⁴
- At least two prescriptions for HRT (in women) in the six months before baseline¹³
- Menopausal symptoms (in women), including vaginal dryness or hot flushes (binary variable yes/no), recorded at baseline.

We restricted all values of these variables to those that had been recorded in the person's electronic healthcare record before baseline, except for BMI and alcohol and smoking status for which we used the values recorded closest to study entry date and recorded before the diagnosis of osteoporotic fracture (or before censoring for those who did not develop a fracture). We assumed that if there was no recorded value of a diagnosis, prescription, or family history then the patient did not have that exposure.

Model derivation and development

We calculated crude incidence rates of osteoporotic fracture (hip, vertebral, or distal radius fracture) and hip fracture by age and sex in the derivation and validation cohorts. We used Cox's proportional hazards models in the derivation dataset to estimate the coefficients and hazard ratios associated with each potential risk factor for the first ever recorded diagnosis of overall fracture and hip fracture for men and women separately. We compared models using the Akaike information criterion (AIC) and the Bayes information criterion (BIC),²⁷ which are likelihood measures in

which lower values indicate better fit and in which a penalty is paid for increasing the number of variables in the model. We used fractional polynomials to model non-linear risk relations with continuous variables where appropriate.²⁸ We tested for interactions between age and smoking; age and parental history of osteoporosis; age and BMI; age and falls; age and use of HRT; use of HRT and smoking; and use of HRT and deprivation. We included significant interactions in the final model when they improved the model fit based on the AIC. Continuous variables were centred for analysis. We checked the assumptions of the proportional hazards model for each variable graphically using log-log survival plots.

After conducting a complete case analysis, we used multiple imputation to replace missing values for alcohol, smoking status, and BMI, and used these values in our main analyses.²⁹⁻³² We used the ICE procedure in Stata³³ to obtain five imputed datasets. Our final model was fitted based on multiply imputed datasets using Rubin's rules to combine effect estimates and estimate standard errors to allow for the uncertainty because of missing data. Multiple imputation is a statistical technique designed to reduce the biases that can occur in "complete case" analysis along with a substantial loss of power and precision.^{29-32,34} The imputation technique involves creating multiple copies of the data and replaces missing values with imputed values based on a suitable random sample from their predicted distribution. Multiple imputation therefore allows patients with incomplete data to still be included in analyses, thereby making full use of all the available data, thus increasing power and precision but without compromising validity.³⁵

We took the regression coefficient (that is, the log of the hazard ratio) for each variable from the final model using multiply imputed data and used these as weights for the QFractureScores. As in previous studies,^{10,36} we combined these weights with the baseline survivor function for diagnosis of fracture or hip fracture obtained from the Cox model evaluated at 10 years and centred on the means of continuous risk factors to derive a risk equation for 10 years' follow-up.

In women we determined the hazard ratios for osteoporotic fracture overall and for hip fracture by HRT use at baseline categorised by (unopposed, cyclical, or continuous) oestrogen dose (high *v* low dose) and type of oestrogen (equine *v* non-equine). These results were incorporated in the QFractureScores for women.

In a separate analysis, we used a time varying Cox regression analysis to examine the effects of duration of use of HRT and time since stopping HRT on risk of fracture in women, treating these terms as time varying covariates. For duration of use of HRT we analysed non-users and new users to determine the hazard ratio of each fracture outcome within one year, two to four years, five to nine years, and 10 years or more of taking HRT compared with no HRT use. We also determined change in risks after stopping HRT categorised as within a year of stopping, one to two years, two to five years, and five or more years. The date of

Table 1 | Characteristics of patients aged 30-85 free of fracture at baseline in derivation and validation cohorts 1993-2008. Figures are numbers (percentage) unless stated otherwise

	Derivation cohort		Validation cohort	
	Women (n=1 183 663)	Men (n=1 174 232)	Women (n=642 153)	Men (n=633 764)
Mean (SD) Townsend score	-0.31 (3.47)	-0.20 (-3.55)	-0.32 (3.33)	-0.23 (3.39)
Median (IQR) age	48 (37-62)	46 (37-59)	49 (37-63)	46 (37-69)
Alcohol consumption:				
Recorded	801 600 (67.72)	717 616 (61.60)	435 452 (67.81)	391 290 (61.74)
Non-drinker	275 984 (23.32)	140 925 (12.00)	148 646 (23.15)	74 718 (11.79)
Trivial <1unit/day	341 295 (28.83)	226 118 (19.26)	185 570 (28.90)	120 989 (19.09)
Light 1-2 units/day	162 433 (13.72)	234 460 (19.97)	89 435 (13.93)	130 813 (20.64)
Moderate 3-6 units/day	19 455 (1.64)	96 202 (8.19)	10 610 (1.65)	54 239 (8.56)
Heavy 7-9 units/day	1208 (0.10)	11 006 (0.94)	618 (0.10)	6005 (0.95)
Very heavy >9 units/day	1231 (0.10)	8877 (0.76)	616 (0.10)	4567 (0.72)
BMI recorded	884 523 (74.73)	781 619 (66.56)	482 194 (75.09)	431 994 (68.16)
Mean (SD) BMI	25.88 (4.86)	26.43 (4.08)	25.82 (4.85)	26.41 (4.02)
Smoking:				
Recorded	1 007 963 (85.16)	929 457 (78.28)	547 531 (85.26)	502 739 (79.33)
Non-smoker	630 470 (53.26)	462 344 (38.94)	340 811 (53.07)	250 715 (39.56)
Ex-smoker	139 496 (11.79)	173 503 (14.61)	75 629 (11.78)	95 004 (14.99)
Current smoker: light	51 945 (4.39)	69 504 (5.85)	29 288 (4.56)	38 173 (6.02)
Current smoker: moderate	131 563 (11.11)	146 959 (12.38)	71 638 (11.16)	76 908 (12.14)
Current smoker: heavy	54 489 (4.60)	77 147 (6.50)	30 165 (4.70)	41 939 (6.62)
Rheumatoid arthritis	9 459 (0.80)	3 903 (0.33)	5 013 (0.78)	2114 (0.33)
Cardiovascular disease	41 842 (3.53)	62 265 (5.24)	23 375 (3.64)	33 542 (5.29)
Type 2 diabetes	22 645 (1.91)	27 637 (2.33)	11 919 (1.86)	14 257 (2.25)
Asthma	66 892 (5.65)	55 888 (4.71)	35 081 (5.46)	28 992 (4.57)
Current tricyclic antidepressants	46 054 (3.89)	14 646 (1.23)	23 729 (3.70)	7354 (1.16)
Current corticosteroids	20 005 (1.69)	11 569 (0.97)	10 509 (1.64)	5792 (0.91)
History of falls	8801 (0.74)	4676 (0.39)	2180 (0.34)	3036 (0.48)
Chronic liver disease	1563 (0.13)	2133 (0.18)	809 (0.13)	1205 (0.19)
Gastrointestinal malabsorption	5970 (0.50)	4851 (0.41)	3346 (0.52)	2595 (0.41)
Other endocrine conditions	8615 (0.73)	1886 (0.16)	5039 (0.78)	1044 (0.16)
Menopausal symptoms	25 683 (2.17)	NA	11 830 (1.84)	NA
Parental history of osteoporosis	5831 (0.49)	441 (0.04)	2180 (0.34)	128 (0.02)

IQR=interquartile range; BMI=body mass index; NA= not applicable.

stopping HRT was taken to be 270 days after the date of the last recorded prescription.

Validation of the QFractureScores

We tested the performances of the final models (QFractureScores) in the validation dataset. We calculated the 10 year estimated risk of sustaining a fracture or hip fracture for each patient in the validation dataset using multiple imputation to replace missing values for alcohol, smoking status, and BMI, as in the derivation dataset.

We calculated the mean predicted fracture risk and the observed fracture risk at 10 years³⁶ and compared these by 10th of predicted risk. The observed risk at 10 years was obtained by using the 10 year Kaplan-Meier estimate. We calculated the D statistic (a measure of discrimination where higher values indicate better discrimination)³⁷ and an R² statistic (which is a measure of explained variation for survival data, where higher values indicate more variation is explained).³⁸ We also calculated the area under the receiver operating characteristics (ROC) curve at 10 years, where higher values indicate better discrimination.

Validation against FRAX (fracture risk assessment)

We compared the performance of the QFractureScore in predicting risk of hip fracture with the performance of the FRAX algorithm using the above validation statistics. FRAX is a relatively new algorithm that predicts 10 year absolute risk of hip fracture and osteoporotic fracture.⁷ It is not currently in widespread use in primary care in the United Kingdom. We used the version that does not incorporate bone mineral density. This version of FRAX is based on the following variables;

- Age
- Sex
- Height
- Weight
- Previous fracture
- Parental history of hip fracture
- Current smoking
- Glucocorticoid treatment
- Rheumatoid arthritis
- Secondary osteoporosis
- Use of alcohol (>3 units/day).

We restricted our comparative analysis to the hip fracture outcome as this is directly comparable between both scores, whereas the FRAX fracture outcome also includes humerus fractures. We used the UK version of the score from the FRAX website (www.shef.ac.uk/FRAX/index.htm) to calculate the 10 year predicted risk of hip fracture for all patients aged 40-85 in the validation dataset, based on relevant input variables of age, sex, height, weight, parent had fractured hip (yes/no), current smoking (yes/no), glucocorticoids (yes/no), rheumatoid arthritis (yes/no), secondary osteoporosis (yes/no), and alcohol >3 units/day (yes/no). Secondary osteoporosis was defined as having liver disease, malabsorption, or endocrine disorders. In all cases previous fracture was counted as negative as we restricted our cohort to patients without a previous fracture. We used the same multiply imputed data that replaced missing values for alcohol use, smoking status, and BMI to calculate the FRAX scores as we used in the validation for the QFractureScores. As with the QFractureScores we assumed that if there was no recorded value of a diagnosis, prescription, or family history then the patient did not have that exposure. We entered the variables for each patient twice, in random order, using automated software to test the reproducibility of the scores generated by the FRAX website.

As we used all the available data on the QResearch database we did not calculate sample size before the

study. All analyses were conducted with Stata (version 10). We chose a significance level of 0.01 (two tailed) as we were considering several variables as potential risk factors in a large dataset and wanted to reduce the risk of having an overly complex model including variables with limited prognostic value.

RESULTS

Description of the derivation and validation dataset

Overall, 535 practices in England and Wales met our inclusion criteria, of which 357 were randomly assigned to the derivation dataset and 178 to the validation dataset.

In the derivation cohort there were 1 204 222 women (1 187 354 men) aged 30-85 at baseline, of whom 20 559 (13 122) had a recorded fracture before the start of the study and were therefore excluded, leaving 1 183 663 (1 174 232) free of fracture at baseline for analysis.

In the validation cohort there were 653 789 women (640 943 men) aged 30-85 at baseline, of whom 11 636 (7179) had a fracture before the start of the study and were therefore excluded, leaving 642 153 (633 764) free of fracture at baseline for analysis.

Table 1 compares the key characteristics of eligible patients in each cohort. While this validation cohort was drawn from an independent group of practices, the baseline characteristics were similar to those for

Table 2 | Incidence rates of osteoporotic fracture (distal radius, hip, or vertebral) and hip fracture per 1000 person years in derivation and validation cohorts by age at baseline in men and women

Age (years)	Derivation cohort				Validation cohort			
	Osteoporotic fractures		Hip fractures		Osteoporotic fractures		Hip fractures	
	No	Rate/1000 (95% CI)	No	Rate/1000 (95% CI)	No	Rate/1000 (95% CI)	No	Rate/1000 (95% CI)
Women								
Total	24 350	3.08 (3.04 to 3.12)	9302	1.15 (1.13 to 1.17)	13 952	3.17 (3.11 to 3.22)	5424	1.2 (1.17 to 1.23)
30-34	605	0.50 (0.46 to 0.54)	23	0.02 (0.01 to 0.03)	322	0.48 (0.43 to 0.54)	8	0.01 (0.01 to 0.02)
35-39	639	0.62 (0.57 to 0.66)	37	0.04 (0.03 to 0.05)	356	0.63 (0.57 to 0.70)	22	0.04 (0.03 to 0.06)
40-44	720	0.88 (0.82 to 0.95)	72	0.09 (0.07 to 0.11)	377	0.83 (0.75 to 0.92)	32	0.07 (0.05 to 0.10)
45-49	1145	1.32 (1.25 to 1.40)	140	0.16 (0.14 to 0.19)	637	1.31 (1.21 to 1.42)	70	0.14 (0.11 to 0.18)
50-54	1745	1.97 (1.88 to 2.07)	229	0.25 (0.22 to 0.29)	941	1.89 (1.77 to 2.01)	129	0.25 (0.21 to 0.30)
55-59	1951	2.70 (2.58 to 2.82)	326	0.44 (0.39 to 0.49)	1156	2.83 (2.67 to 3.00)	220	0.53 (0.46 to 0.60)
60-64	2578	3.99 (3.84 to 4.15)	615	0.92 (0.85 to 1.00)	1451	4.03 (3.83 to 4.24)	407	1.09 (0.99 to 1.20)
65-69	3282	5.72 (5.53 to 5.92)	1152	1.92 (1.81 to 2.04)	1888	5.81 (5.56 to 6.08)	677	1.99 (1.85 to 2.15)
70-74	4086	8.05 (7.81 to 8.30)	1899	3.55 (3.39 to 3.71)	2378	8.35 (8.02 to 8.69)	1085	3.59 (3.39 to 3.81)
75-85	7599	12.11 (11.84 to 12.38)	4809	7.19 (6.99 to 7.40)	4446	12.39 (12.04 to 12.76)	2774	7.23 (6.97 to 7.51)
Men								
Total	7934	0.99 (0.96 to 1.01)	3067	0.38 (0.36 to 0.39)	4519	1.01 (0.98 to 1.04)	1738	0.38 (0.37 to 0.40)
30-34	693	0.54 (0.50 to 0.58)	46	0.04 (0.03 to 0.05)	344	0.49 (0.44 to 0.54)	26	0.04 (0.02 to 0.05)
35-39	677	0.57 (0.53 to 0.61)	58	0.05 (0.04 to 0.06)	354	0.54 (0.48 to 0.60)	38	0.06 (0.04 to 0.08)
40-44	592	0.57 (0.53 to 0.62)	88	0.08 (0.07 to 0.10)	339	0.58 (0.52 to 0.64)	40	0.07 (0.05 to 0.09)
45-49	610	0.61 (0.56 to 0.66)	109	0.11 (0.09 to 0.13)	353	0.63 (0.56 to 0.70)	60	0.11 (0.08 to 0.14)
50-54	629	0.72 (0.67 to 0.78)	135	0.15 (0.13 to 0.18)	354	0.73 (0.66 to 0.81)	76	0.16 (0.12 to 0.19)
55-59	629	0.87 (0.80 to 0.94)	189	0.26 (0.22 to 0.30)	381	0.94 (0.85 to 1.04)	104	0.25 (0.21 to 0.31)
60-64	661	1.06 (0.98 to 1.14)	251	0.4 (0.35 to 0.45)	375	1.09 (0.98 to 1.20)	148	0.42 (0.36 to 0.50)
65-69	787	1.49 (1.39 to 1.60)	398	0.75 (0.68 to 0.82)	461	1.60 (1.46 to 1.75)	237	0.81 (0.71 to 0.92)
70-74	1014	2.54 (2.39 to 2.70)	599	1.48 (1.37 to 1.61)	603	2.72 (2.51 to 2.94)	355	1.58 (1.42 to 1.75)
75-85	1642	4.35 (4.15 to 4.57)	1194	3.13 (2.95 to 3.31)	955	4.51 (4.23 to 4.80)	654	3.05 (2.82 to 3.29)

Table 3 | Characteristics of patients with and without body mass index (BMI) and with and without smoking status recorded. Figures are numbers (percentage) unless stated otherwise

	Women		Men	
According to BMI recording				
BMI recorded	Yes	No	Yes	No
Total	870 874	312 789	781 619	392 613
Mean (SD) age (years)	49.34 (14.17)	54.87 (17.70)	48.15 (13.41)	50.38 (15.87)
Mean (SD) Townsend	−0.45 (3.43)	0.06 (3.55)	−0.35 (3.51)	0.11 (3.62)
Current smoker	210 541 (24.18)	39 090 (12.50)	246 484 (31.54)	56 324 (14.35)
Parental history of osteoporosis	5377 (0.62)	454 (0.15)	396 (0.05)	45 (0.01)
Rheumatoid arthritis	6703 (0.77)	2756 (0.88)	2605 (0.33)	1298 (0.33)
Cardiovascular disease	6575 (0.75)	35 267 (11.28)	7616 (0.97)	54 649 (13.92)
Type 2 diabetes	16 589 (1.90)	6056 (1.94)	19 732 (2.52)	7905 (2.01)
Asthma	55 049 (6.32)	11 843 (3.79)	43 678 (5.59)	12 210 (3.11)
Current tricyclic antidepressants	35 835 (4.11)	10 219 (3.27)	10 272 (1.31)	4374 (1.11)
Current corticosteroids	15 139 (1.74)	4866 (1.56)	8042 (1.03)	3527 (0.90)
History of falls	5741 (0.66)	3060 (0.98)	2809 (0.36)	1867 (0.48)
Chronic liver disease	1090 (0.13)	473 (0.15)	1424 (0.18)	709 (0.18)
Gastrointestinal malabsorption	4814 (0.55)	1156 (0.37)	3582 (0.46)	1269 (0.32)
Other endocrine conditions	6549 (0.75)	2066 (0.66)	1336 (0.17)	550 (0.14)
Menopausal symptoms	21 889 (2.51)	3794 (1.21)	NA	NA
10 year osteoporotic fracture risk	2.62	4.30	0.9	1.12
10 year hip fracture risk	0.81	2.19	0.29	0.56
According to smoking recording				
Smoking recorded	Yes	No	Yes	No
Total	1 007 963	175 700	929 457	244 775
Mean (SD) age (years)	49.67 (14.53)	57.3 (18.26)	48.24 (13.62)	51.39 (16.49)
Mean (SD) Townsend	−0.41 (3.45)	0.23 (3.57)	−0.32 (3.52)	0.29 (3.63)
Current smoker	249 631 (24.77)	0 (0.00)	302 808 (32.58)	0 (0.00)
Parental history of osteoporosis	5719 (0.57)	112 (0.06)	422 (0.05)	19 (0.01)
Rheumatoid arthritis	8021 (0.80)	1438 (0.82)	3136 (0.34)	767 (0.31)
Cardiovascular disease	11 437 (1.13)	30 405 (17.31)	18 525 (1.99)	43 740 (17.87)
Type 2 diabetes	18 359 (1.82)	4286 (2.44)	21733 (2.34)	5904 (2.41)
Asthma	61 741 (6.13)	5151 (2.93)	49 987 (5.38)	5901 (2.41)
Current tricyclic antidepressants	40 834 (4.05)	5220 (2.97)	12 169 (1.31)	2477 (1.01)
Current corticosteroids	17 476 (1.73)	2529 (1.44)	9683 (1.04)	1886 (0.77)
History of falls	7104 (0.70)	1697 (0.97)	3621 (0.39)	1055 (0.43)
Chronic liver disease	1310 (0.13)	253 (0.14)	1715 (0.18)	418 (0.17)
Gastrointestinal malabsorption	5428 (0.54)	542 (0.31)	4201 (0.45)	650 (0.27)
Other endocrine conditions	7507 (0.74)	1108 (0.63)	1542 (0.17)	344 (0.14)
Menopausal symptoms	24 226 (2.40)	1457 (0.83)	NA	NA
10 year osteoporotic fracture risk	2.78	4.88	0.94	1.07
10 year hip fracture risk	0.93	2.71	0.58	0.33
NA=not applicable.				

NA=not applicable.

the derivation cohort across all measures in both men and women.

Table 2 shows the incidence rates in each cohort. During the 7 898 208 person years of follow-up for women in the derivation cohort 24 350 fractures were recorded (hip, vertebral, or distal radius), giving an overall incidence rate of 3.08 per 1000 person years (95% confidence interval 3.04 to 3.12). For men, there were 7934 incident fractures arising from 8 049 306 person years, giving an incidence rate of 0.99 per 1000 person years (0.96 to 1.01). In women, 38.2% of the fractures were hip fractures, in men the corresponding figure was 38.9%. Similar incidence

rates were found in the validation cohort (table 2). Incidence rates were higher in women than in men and rose steeply with age. In the derivation cohort, highest fracture rates were observed among those aged 75 and over at baseline: the incidence rate was 12.11 per 1000 person years (11.84 to 12.38) in women and 4.35 per 1000 person years (4.15 to 4.57) in men. The corresponding figures for incidence of hip fracture among those aged 75 and over were 7.19 per 1000 person years (6.99 to 7.40) in women and 3.13 per 1000 person years (2.95 to 3.31) in men.

Tables 3 and 4 show the characteristics of patients with and without BMI, smoking status, and alcohol

status recorded. Table 4 also shows the characteristics of patients with complete data for all three variables. There were differences in observed characteristics between those with and without missing data, supporting the assumption that data are missing at random, which supports the use of multiple imputation.

Model development

Table 5 shows the results of the multivariate final Cox regression analysis for fracture and hip fracture in men based on a complete cases analysis and using multiply imputed data. Table 6 shows the results for women.

There was no evidence that the proportional hazards assumption was not valid in any of the models presented.

Risk factors for fracture in men

After adjustment for all other variables in the model, we found significant associations with overall risk of fracture and risk of hip fracture in men for the following variables, which were therefore included in both final algorithms for men: age, BMI, smoking status, alcohol use, rheumatoid arthritis, cardiovascular disease, type 2 diabetes, asthma, use of tricyclic antidepressants, history of falls, and liver disease. For

Table 4 | Characteristics of patients with and without alcohol status recorded and with and without complete data for body mass index (BMI), smoking status, and alcohol status. Figures are numbers (percentage) unless stated otherwise

	Women		Men	
According to alcohol recording				
Alcohol recorded	Yes	No	Yes	No
Total	801 606	382 057	717 588	456 644
Mean (SD) age (years)	49.63 (14.20)	53.26 (17.35)	48.25 (13.30)	49.91 (15.74)
Mean (SD) Townsend	-0.46 (3.43)	-0.01 (3.54)	-0.39 (3.50)	0.11 (3.61)
Current smoker	194 143 (24.22)	55 488 (14.52)	227 718 (31.73)	75 090 (16.44)
Parental history of osteoporosis	5284 (0.66)	547 (0.14)	400 (0.06)	41 (0.01)
Rheumatoid arthritis	6544 (0.82)	2915 (0.76)	2516 (0.35)	1387 (0.30)
Cardiovascular disease	6439 (0.80)	35 403 (9.27)	7545 (1.05)	54 720 (11.98)
Type 2 diabetes	15 618 (1.95)	7027 (1.84)	18 201 (2.54)	9436 (2.07)
Asthma	50 654 (6.32)	16 238 (4.25)	39 318 (5.48)	16 570 (3.63)
Current tricyclic antidepressants	33 865 (4.22)	12 189 (3.19)	9659 (1.35)	4987 (1.09)
Current corticosteroids	13 909 (1.74)	6096 (1.60)	7236 (1.01)	4333 (0.95)
History of falls	5765 (0.72)	3036 (0.79)	2858 (0.40)	1818 (0.40)
Chronic liver disease	1051 (0.13)	512 (0.13)	1410 (0.20)	723 (0.16)
Gastrointestinal malabsorption	4506 (0.56)	1464 (0.38)	3395 (0.47)	1456 (0.32)
Other endocrine conditions	6231 (0.78)	2384 (0.62)	1275 (0.18)	611 (0.13)
Menopausal symptoms	21 159 (2.64)	4524 (1.18)	NA	NA
10 year osteoporotic fracture risk	2.75	3.68	0.95	0.98
10 year hip fracture risk	0.88	1.79	0.32	0.47
According to completeness of data				
Complete data	Yes	No	Yes	No
Total	732 931	450 732	649 500	524 732
Mean (SD) age (years)	49.42 (13.94)	53.06 (17.24)	48.27 (13.17)	49.67 (15.59)
Mean (SD) Townsend	-0.5 (3.42)	-0.01 (3.54)	-0.42 (3.49)	0.08 (3.61)
10 year fracture risk	0.9737	0.9621	0.9908	0.9897
Current smoker	177 512 (24.22)	72 119 (16.00)	204 667 (31.51)	98 141 (18.70)
Parental history of osteoporosis	5071 (0.69)	760 (0.17)	387 (0.06)	54 (0.01)
Rheumatoid arthritis	5848 (0.80)	3611 (0.80)	2261 (0.35)	1642 (0.31)
Cardiovascular disease	4802 (0.66)	37 040 (8.22)	5433 (0.84)	56 832 (10.83)
Type 2 diabetes	14 492 (1.98)	8153 (1.81)	17 237 (2.65)	10 400 (1.98)
Asthma	47 202 (6.44)	19 690 (4.37)	36 652 (5.64)	19 236 (3.67)
Current tricyclic antidepressants	31 111 (4.24)	14 943 (3.32)	8721 (1.34)	5925 (1.13)
Current corticosteroids	12 743 (1.74)	7262 (1.61)	6547 (1.01)	5022 (0.96)
History of falls	5077 (0.69)	3724 (0.83)	2467 (0.38)	2209 (0.42)
Chronic liver disease	915 (0.12)	648 (0.14)	1221 (0.19)	912 (0.17)
Gastrointestinal malabsorption	4197 (0.57)	1773 (0.39)	3120 (0.48)	1731 (0.33)
Other endocrine conditions	5741 (0.78)	2874 (0.64)	1180 (0.18)	706 (0.13)
Menopausal symptoms	19 991 (2.73)	5692 (1.26)	NA	NA
10 year osteoporotic fracture risk	2.63	2.63	0.92	1.03
10 year hip fracture risk	0.79	1.83	0.29	0.50
NA=not applicable.				

NA=not applicable.

Table 5 Adjusted hazard ratios* (95% confidence intervals) for osteoporotic fracture (distal radius, hip, or vertebral) and hip fracture in men

	Osteoporotic fracture		Hip fracture	
	Complete case analysis	Multiply imputed data	Complete case analysis	Multiply imputed data
Non-smoker	1.00	1.00	1.00	1.00
Ex-smoker	0.98 (0.91 to 1.05)	0.98 (0.92 to 1.05)	0.99 (0.87 to 1.11)	0.99 (0.89 to 1.10)
Current smoker:				
Light	1.06 (0.95 to 1.17)	1.13 (1.03 to 1.23)	1.23 (1.03 to 1.46)	1.21 (1.05 to 1.40)
Moderate	1.24 (1.13 to 1.36)	1.20 (1.12 to 1.29)	1.61 (1.37 to 1.91)	1.47 (1.26 to 1.71)
Heavy	1.4 (1.27 to 1.55)	1.36 (1.25 to 1.47)	2.18 (1.82 to 2.62)	1.70 (1.44 to 2.00)
Alcohol:				
Non-drinker	1.00	1.00	1.00	1.00
Trivial <1unit/day	0.91 (0.84 to 0.99)	0.93 (0.86 to 0.99)	0.81 (0.71 to 0.92)	0.78 (0.71 to 0.87)
Light 1-2 units/day	0.95 (0.88 to 1.03)	0.95 (0.89 to 1.02)	0.84 (0.74 to 0.97)	0.77 (0.70 to 0.86)
Moderate 3-6 units/day	1.08 (0.98 to 1.19)	1.05 (0.95 to 1.16)	0.91 (0.76 to 1.09)	0.81 (0.70 to 0.93)
Heavy 7-9 units/day	1.06 (0.83 to 1.35)	1.10 (0.89 to 1.36)	0.85 (0.52 to 1.38)	0.82 (0.56 to 1.20)
Very heavy >9 units/day	1.84 (1.47 to 2.29)	1.62 (1.34 to 1.96)	2.56 (1.78 to 3.67)	1.70 (1.20 to 2.43)
Rheumatoid arthritis†	1.41 (1.01 to 1.97)	1.48 (1.15 to 1.90)	1.81 (1.15 to 2.85)	1.81 (1.31 to 2.50)
Cardiovascular disease†	1.11 (0.89 to 1.39)	1.29 (1.20 to 1.39)	1.15 (0.84 to 1.58)	1.24 (1.13 to 1.37)
Type 2 diabetes†	1.18 (1.02 to 1.37)	1.20 (1.07 to 1.35)	1.42 (1.15 to 1.74)	1.38 (1.18 to 1.62)
Asthma†	1.24 (1.10 to 1.39)	1.28 (1.17 to 1.41)	1.24 (1.01 to 1.52)	1.31 (1.12 to 1.53)
Current tricyclic antidepressants†	1.40 (1.18 to 1.67)	1.36 (1.18 to 1.56)	1.77 (1.37 to 2.28)	1.67 (1.38 to 2.01)
Current corticosteroids†	1.65 (1.39 to 1.97)	1.46 (1.27 to 1.68)	1.61 (1.23 to 2.10)	1.22 (0.99 to 1.51)
History of falls†	2.17 (1.60 to 2.93)	2.23 (1.80 to 2.75)	2.29 (1.46 to 3.61)	2.66 (2.03 to 3.49)
Liver disease†	3.59 (2.45 to 5.24)	2.86 (2.04 to 4.01)	3.75 (2.01 to 6.99)	2.96 (1.75 to 5.01)

BMI=body mass index.

*Hazard ratios simultaneously adjusted for all other variables shown in table as well as fractional polynomial terms for age and BMI. Fractional polynomial terms for age and BMI were: (age/10) and (age/10)² and (BMI/10)⁻² for osteoporotic fracture; and (age/10)² and log(BMI/10) and (log(BMI/10))² for hip fracture.

†Compared with patients without condition/medication at baseline.

consistency, we also included current use of corticosteroids in the final models for osteoporotic fracture and hip fracture, although it tended towards significance only for hip fracture ($P=0.067$), which might reflect the lower numbers of patients with a hip fracture. Variables that were not significant on multivariate analyses in men (and that were not therefore included in the final algorithms) included deprivation, gastrointestinal malabsorption, other endocrine conditions, and parental history of osteoporosis. There were also no significant interactions.

Table 5 shows the adjusted hazard ratios for the variables included in both final algorithms for men based on the multiply imputed data; fractional polynomial terms for age and BMI were also included in the algorithms. Patients with liver disease had a 196% increased risk of hip fracture after adjustment for all other variables. Similarly, patients with a history of falls had a 166% increased risk of hip fracture. Heavy smokers had a 70% increased risk of hip fracture compared with non-smokers; patients with very heavy alcohol intake had a 70% increased risk compared with non-drinkers. Compared with patients without each disease, patients with rheumatoid arthritis had an 81% increased risk of hip fracture; those with cardiovascular disease had a 24% increased risk; those with type 2 diabetes had a 38% increased risk; and those with asthma had a 31% increased risk. Patients taking

tricyclic antidepressants had a 67% increased risk of hip fracture, and those prescribed steroids had a 22% increased risk.

Figure 1 shows the estimated adjusted hazard ratios with the fractional polynomial terms for age and BMI in men. There were two age terms for the osteoporotic fracture outcome (age/10) and (age/10)² and one term for BMI (bmi/10)⁻². There was a single age term for hip fracture, which was (age/10)² with two terms for BMI: log(BMI/10) and (log(BMI/10))².

Risk factors for fracture in women

After adjustment for all other variables in the model, we found significant associations with overall fracture risk in women for the following variables: use of HRT, smoking status, use of alcohol, parental history of osteoporosis, rheumatoid arthritis, cardiovascular disease, type 2 diabetes, asthma, tricyclic antidepressants, use of corticosteroids, history of falls, menopausal symptoms, chronic liver disease, gastrointestinal malabsorption, and other endocrine disorders (table 6). There were also significant associations with age and BMI with fractional polynomial terms. The final algorithm for osteoporotic fracture in women included all of these variables. Table 6 shows the adjusted hazard ratios.

Some variables were significantly associated with overall risk of fracture but not with risk of hip fracture at the 0.01 level. These were use of HRT, menopausal

symptoms, parental history of osteoporosis, malabsorption, and other endocrine disorders. The magnitude and direction of the coefficients were similar to those for overall risk of fracture so they were included in the final hip fracture model for consistency. The final algorithm for risk of hip fracture in women included age, BMI, smoking status, alcohol use, use of HRT, parental history of osteoporosis, rheumatoid arthritis, cardiovascular disease, type 2 diabetes, asthma, current use of tricyclic antidepressants, current use of corticosteroids, history of falls, menopausal symptoms, liver disease, gastrointestinal malabsorption, and other endocrine disorders. Table 6 shows the adjusted hazard ratios for variables included in both final algorithms for women based on the multiply imputed data.

Figure 1 shows the estimated adjusted hazard ratios with the fractional polynomial terms for age and BMI in women. There were two terms for age for the fracture outcome, these were $(\text{age}/10)^{-1}$ and $(\text{age}/10)^{-1}\log(\text{age}/10)$. There was one fractional polynomial term for BMI, which was $(\text{BMI}/10)^{-1}$. The terms selected for inclusion in the hip fracture model in women were $(\text{age}/10)^3$ and $(\text{age}/10)^3\log(\text{age}/10)$ and one term for BMI, which was $(\text{BMI}/10)^{-2}$.

There were significant interactions between age and BMI and between age and parental history of osteoporosis for overall fracture in women. The more complex model including the age interaction terms, however, did not improve the model fit statistics and resulted in similar predicted scores compared with the simpler model so we selected the more parsimonious version as our final model.

Effect of hormone replacement therapy on fracture risk

Overall, 168 536 women (14.24% of 1 183 663) were prescribed HRT at baseline. Of these, 16 425 (9.75%) were prescribed low dose oestrogen equine; 30 598 (18.16%) were prescribed low dose non-equine; 7474 (4.43%) were prescribed high dose equine; 7205 (4.28%) were prescribed high dose non-equine; 23 430 (13.90%) were prescribed cyclical low dose equine; 28 427 (16.87%) were prescribed cyclical low dose non-equine; 2342 (1.392%) were prescribed cyclical high dose equine; 16 753 (9.94%) were prescribed cyclical high dose non-equine; 6765 (4.01%) were prescribed continuous low dose equine; 11 629 (6.90%) were prescribed continuous low dose non-equine; 14 186 (8.42%) were prescribed continuous high dose non-equine; and 3302 (1.96%) were prescribed tibolone.

We found significant associations between risk of fracture and some types of HRT (table 6). Women prescribed unopposed oestrogen HRT had a reduced risk of osteoporotic fracture. There was a 25% decrease (8% to 38%) in risk of fracture with high dose unopposed equine oestrogen, and a 19% decreased risk (5% to 30%) for cyclical high dose non-equine, both of which were significant at the 0.01 level. There was a borderline significant decrease in risk ($0.01 < P < 0.05$) with low dose unopposed equine oestrogen, low dose unopposed non-equine oestrogen, high dose unopposed non-equine oestrogen, and continuous high dose non-equine. There was a 24% increase in risk with continuous low dose equine HRT (6% to 45%). The direction and the magnitude of the risks associated with HRT use for the hip fracture outcome were similar, although they were not significant at the 0.01 level.

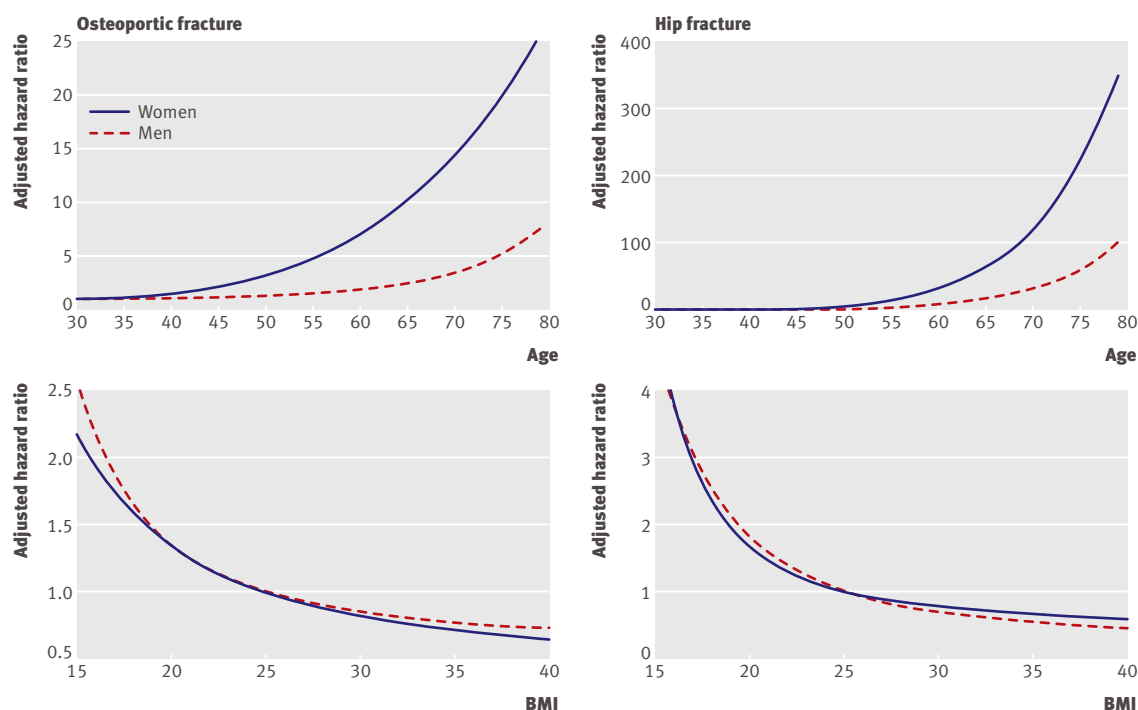


Fig 1 | Adjusted hazard ratios for age (compared with age 30) and BMI (compared with BMI=25) for osteoporotic fracture (distal radius, hip, or vertebral) and for hip fracture in men and women with fractional polynomial terms

Table 6 | Adjusted hazard ratios* (95% confidence intervals) for osteoporotic fractures (distal radius, hip, vertebral) and hip fracture in women

	Osteoporotic fracture		Hip fracture	
	Complete case analysis	Multiply imputed data	Complete case analysis	Multiply imputed data
No use of HRT	1.00	1.00	1.00	1.00
Type of HRT†:				
Low dose unopposed equine oestrogen	0.91 (0.81 to 1.02)	0.90 (0.81 to 0.99)	0.84 (0.66 to 1.06)	0.80 (0.66 to 0.97)
Low dose unopposed non-equine oestrogen	0.90 (0.82 to 0.99)	0.90 (0.82 to 0.98)	0.96 (0.79 to 1.17)	0.87 (0.74 to 1.03)
High dose unopposed equine oestrogen	0.72 (0.57 to 0.90)	0.75 (0.62 to 0.92)	0.66 (0.37 to 1.16)	0.65 (0.40 to 1.06)
High dose unopposed non-equine oestrogen	0.73 (0.57 to 0.94)	0.76 (0.61 to 0.96)	0.74 (0.40 to 1.38)	0.80 (0.48 to 1.33)
Cyclical low dose equine	0.90 (0.81 to 1.01)	0.92 (0.83 to 1.02)	0.91 (0.68 to 1.21)	0.93 (0.72 to 1.19)
Cyclical low dose non-equine	0.88 (0.78 to 0.99)	0.94 (0.84 to 1.04)	0.78 (0.56 to 1.08)	0.80 (0.60 to 1.07)
Cyclical high dose equine	1.14 (0.81 to 1.60)	1.14 (0.84 to 1.56)	1.06 (0.44 to 2.56)	1.00 (0.45 to 2.23)
Cyclical high dose non-equine	0.79 (0.67 to 0.93)	0.81 (0.70 to 0.94)	0.93 (0.61 to 1.42)	0.91 (0.62 to 1.33)
Continuous low dose equine	1.16 (0.96 to 1.39)	1.24 (1.06 to 1.45)	1.20 (0.82 to 1.77)	1.12 (0.80 to 1.56)
Continuous low dose non-equine	0.97 (0.83 to 1.14)	1.03 (0.89 to 1.19)	0.87 (0.59 to 1.30)	0.84 (0.59 to 1.20)
Continuous high dose non-equine	0.82 (0.72 to 0.93)	0.88 (0.78 to 0.99)	0.73 (0.54 to 1.00)	0.82 (0.64 to 1.06)
Tibolone	0.86 (0.67 to 1.10)	0.92 (0.74 to 1.14)	0.41 (0.19 to 0.86)	0.62 (0.37 to 1.03)
Non-smoker	1.00	1.00	1.00	1.00
Ex-smoker	1.02 (0.97 to 1.06)	1.02 (0.98 to 1.06)	1.05 (0.97 to 1.14)	1.04 (0.97 to 1.11)
Current smoker:				
Light	1.02 (0.95 to 1.10)	1.04 (0.98 to 1.11)	1.21 (1.06 to 1.37)	1.18 (1.05 to 1.32)
Moderate	1.14 (1.07 to 1.20)	1.11 (1.06 to 1.17)	1.42 (1.28 to 1.58)	1.32 (1.22 to 1.43)
Heavy	1.21 (1.12 to 1.31)	1.21 (1.13 to 1.29)	1.87 (1.62 to 2.16)	1.55 (1.39 to 1.73)
Alcohol:				
Non-drinker	1.00	1.00	1.00	1.00
Trivial <1 units/day	1.00 (0.97 to 1.04)	1.00 (0.97 to 1.03)	0.92 (0.86 to 0.97)	0.88 (0.82 to 0.94)
Light 1-2 units/day	1.04 (0.99 to 1.09)	1.02 (0.97 to 1.07)	0.91 (0.84 to 0.99)	0.87 (0.81 to 0.94)
Moderate 3-6 units/day	1.08 (0.96 to 1.21)	1.03 (0.92 to 1.14)	1.10 (0.89 to 1.35)	0.95 (0.77 to 1.18)
Heavy 7-9 units/day	1.58 (1.06 to 2.36)	1.48 (1.06 to 2.06)	1.60 (0.76 to 3.35)	1.41 (0.81 to 2.45)
Very heavy >9 units/day	2.40 (1.66 to 3.46)	2.01 (1.49 to 2.71)	2.93 (1.57 to 5.47)	2.25 (1.35 to 3.73)
Parental history osteoporosis†	1.63 (1.38 to 1.92)	1.78 (1.54 to 2.07)	1.09 (0.69 to 1.71)	1.06 (0.71 to 1.58)
Rheumatoid arthritis†	1.26 (1.11 to 1.43)	1.31 (1.18 to 1.44)	1.82 (1.52 to 2.18)	1.78 (1.56 to 2.03)
Cardiovascular disease†	1.12 (0.99 to 1.28)	1.17 (1.12 to 1.22)	1.26 (1.04 to 1.52)	1.17 (1.09 to 1.24)
Type 2 diabetes†	1.27 (1.17 to 1.39)	1.25 (1.17 to 1.34)	1.79 (1.59 to 2.02)	1.67 (1.51 to 1.83)
Asthma†	1.28 (1.20 to 1.36)	1.29 (1.22 to 1.36)	1.39 (1.24 to 1.55)	1.32 (1.21 to 1.44)
Current tricyclic antidepressants†	1.29 (1.21 to 1.37)	1.31 (1.25 to 1.38)	1.31 (1.18 to 1.46)	1.34 (1.24 to 1.45)
Current corticosteroids†	1.17 (1.07 to 1.28)	1.17 (1.09 to 1.26)	1.13 (0.97 to 1.31)	1.18 (1.05 to 1.31)
History of falls†	1.65 (1.45 to 1.87)	1.82 (1.66 to 1.99)	1.69 (1.40 to 2.05)	2.03 (1.80 to 2.29)
Menopausal symptoms†	1.13 (1.03 to 1.23)	1.15 (1.07 to 1.24)	1.16 (0.99 to 1.36)	1.13 (0.99 to 1.28)
Liver disease†	1.79 (1.30 to 2.46)	1.79 (1.41 to 2.27)	1.75 (1.02 to 3.02)	1.65 (1.13 to 2.42)
Gastrointestinal malabsorption†	1.32 (1.11 to 1.57)	1.23 (1.06 to 1.43)	1.29 (0.94 to 1.76)	1.10 (0.85 to 1.42)
Other endocrine disorder†	1.10 (0.95 to 1.26)	1.11 (1.00 to 1.25)	1.09 (0.69 to 1.71)	1.19 (1.01 to 1.40)

BMI=body mass index; HRT= hormone replacement therapy.

*Hazard ratios simultaneously adjusted for all variables shown in table as well as fractional polynomial terms for age and BMI. Fractional polynomial terms for age and BMI were: (age/10)⁻¹ and (age/10)⁻¹log(age/10) and (BMI/10)⁻¹ for osteoporotic fracture; and (age/10)³ and (age/10)³log(age/10) and (BMI/10)⁻² for hip fracture.

Low dose unopposed equine oestrogen, however, had a borderline association with a 20% decrease in risk (2% to 34%, $P=0.026$).

Table 7 shows the results of the time varying analyses for duration of use of HRT and time since stopping HRT. There was a trend over time with increasing duration of use of HRT, with no significant effect of HRT on overall risk of fracture in the first year of use, a 36% significant decrease in risk (8% to 40%) for one to two years of use, and a 45% decreased risk among women with 10 or more years of use (16% to 64%). There was no significant increase or decrease in risk within the first

year after stopping HRT, although there was a 23% increase in risk (8% to 42%) one to two years after stopping and a 23% increase (12% to 34%) two to five years after stopping compared with women who had not been prescribed HRT. The pattern for the time varying analysis for the hip fracture outcome with duration of use of HRT was similar: the decrease ranged between a 58% reduction in risk after one to two years of use and a 31% reduction after two to five years. Although the confidence intervals were wider because of the smaller number of incident cases, there was no association with time since stopping HRT

Table 7 | Risk of osteoporotic fractures and hip fracture for duration of use of HRT and time since stopped HRT in women from time varying covariate analysis. Figures are adjusted hazard ratios* (95% confidence intervals)

	Osteoporotic fracture		Hip fracture	
	HR (95% CI)	P value	HR (95% CI)	P value
Not taking HRT	1.00		1.00	
Taking HRT:				
1 year	0.86 (0.73 to 1.02)	0.08	0.73 (0.51 to 1.06)	0.09
1-2 years	0.74 (0.60 to 0.92)	0.006	0.41 (0.22 to 0.80)	0.01
2-5 years	0.70 (0.61 to 0.81)	<0.001	0.69 (0.49 to 0.99)	0.04
5-9 years	0.70 (0.59 to 0.82)	<0.001	0.83 (0.58 to 1.17)	0.28
≥10 years	0.55 (0.36 to 0.84)	0.006	0.73 (0.33 to 1.63)	0.44
Stopped HRT:				
<1 year	1.07 (0.93 to 1.23)	0.36	1.09 (0.81 to 1.46)	0.58
1-2 years	1.23 (1.08 to 1.42)	0.003	0.96 (0.69 to 1.32)	0.78
2-5 years	1.23 (1.12 to 1.34)	<0.001	1.05 (0.87 to 1.28)	0.61
≥5 years	1.12 (1.01 to 1.24)	0.03	0.96 (0.79 to 1.18)	0.69

HRT= hormone replacement therapy.

*Hazard ratios adjusted for variables included in final models for osteoporotic fracture and hip fracture in women as shown in table 3.

Validation of the QFractureScores

Discrimination

Table 8 shows the discrimination statistics for the QFractureScores for men and women for both fracture outcomes. The hip fracture algorithm had the best performance among both men and women. It explained 63.94% (62.12% to 65.76%) of the variation in women and 63.19% (60.81% to 65.57%) of the variation in men. The D statistic values were high for women (2.73, 2.62 to 2.83) and men (2.68, 2.55 to 2.82) and were over twice the magnitude of the corresponding D statistic results for osteoporotic fracture in men and women. The ROC values for hip fracture were high with values of 0.89 for women and 0.86 for men compared with 0.79 and 0.69, respectively, for the overall fracture outcome.

Calibration

Table 9 compares the mean predicted scores applying the QFractureScores with the observed risks at 10 years within each 10th of predicted risk to assess the calibration of the model in the validation sample. There was close correspondence between predicted and observed 10 year risks within each model 10th for overall fracture. For example, in the top 10th of risk, the mean predicted 10 year risk of fracture was 12.9% and the observed risk was 13.0%. The ratio of predicted to observed risk in this 10th was 0.99, indicating almost perfect calibration (a ratio of 1 indicates perfect calibration—that is, no underprediction or overprediction). Similar results were obtained for men with a ratio of 1.0 in the top 10th of predicted risk. For hip fracture there was also close correspondence, except for overprediction in the lowest 10th in both sexes and the third lowest in women. There was also 19% overprediction of risk among the top 10th of predicted risk for hip fracture in men, with 4% overprediction for women.

Table 9 also shows the decile cut-offs for men and women for each of the QFractureScores and the number and proportion of incident cases in each 10th. For women for the fracture outcome, 33.72% of incident cases fell in the top 10th and 31.89% for men. For women for the hip fracture outcome, 52.69% of incident cases fell in the top 10th and 54.20% of cases for men.

Validation of the fracture clinical risk score against FRAX

We calculated a hip fracture score using the FRAX algorithm for 454 499 women aged 40-85 and 424 336 men in the validation cohort. The D statistic for hip fracture for the FRAX algorithm was 2.26 (2.21 to 2.30) for women and 2.22 (2.14 to 2.30) for men. The FRAX algorithm explained 54.83% (54.43% to 55.12%) of the variation in women and 54.07% (52.10% to 53.65%) in men. The ROC value for the FRAX algorithm was 0.845 for women and 0.817 for men.

We recalculated the validation statistics for the QFractureScores restricting the population to patients aged 40-85. The D statistic for hip fracture was 2.37 (2.32 to 2.42) for women and 2.39 (2.30 to 2.48) for men. The QFractureScores algorithm explained 57.29% (57.18% to 58.09%) of the variation in women and 57.67% (56.78% to 58.57%) in men.

Table 10 and figure 2 show the calibration statistics for patients aged 40-85. FRAX tended to overpredict the risk of hip fracture within each 10th of risk, as shown by the ratio of the predicted risk to the observed risk.

DISCUSSION

Summary of main findings

A new risk prediction algorithm (the QFractureScore) for estimating the 10 year absolute risk of osteoporotic fracture and hip fracture in men and women shows some evidence of improved discrimination and calibration compared with the FRAX algorithm. We addressed some of research questions highlighted by NICE guidance,⁶ National Osteoporosis Guidelines,⁵ and the WHO.¹¹ Our algorithm extends the age range and quantifies additional risk factors not fully taken account of in FRAX, such as falls, type 2 diabetes, cardiovascular disease, use of HRT, menopausal symptoms, and use of tricyclic antidepressants. We also validated the QFractureScores algorithms

Table 8 | Validation statistics* for QFractureScore for osteoporotic fracture and hip fracture in validation cohort. Figures are means (95% confidence intervals)

	Osteoporotic fracture	Hip fracture
Women		
R ²	44.87 (43.07 to 46.67)	63.94 (62.12 to 65.76)
D statistic	1.85 (1.78 to 1.91)	2.73 (2.62 to 2.83)
ROC statistic	0.788 (0.786 to 0.790)	0.890 (0.889 to 0.892)
Men		
R ²	30.02 (22.21 to 37.84)	63.19 (60.81 to 65.57)
D statistic	1.34 (1.09 to 1.59)	2.68 (2.55 to 2.82)
ROC statistic	0.688 (0.684 to 0.692)	0.856 (0.851 to 0.860)

ROC=receiver operating characteristics.

*Higher values indicate better discrimination.

Table 9 | Predicted and observed risk for osteoporotic fracture and hip fracture at 10 years by 10th of predicted risk using QFractureScores in patients aged 30-85

	Osteoporotic fracture					Hip fracture				
	Decile cut-offs*	Incident cases (column %)	Mean predicted risk (%)	Observed risk (%)	Ratio predicted/observed	Decile cut-offs*	Incident cases (column %)	Mean predicted risk (%)	Observed risk (%)	Ratio predicted/observed
Women										
1	—	176 (1.26)	0.41	0.42	0.96		5 (0.09)	0.02	0.01	1.86
2	0.47	223 (1.6)	0.52	0.57	0.92	0.02	11 (0.2)	0.03	0.03	0.94
3	0.57	246 (1.76)	0.63	0.58	1.09	0.03	9 (0.17)	0.04	0.02	2.02
4	0.70	355 (2.54)	0.81	0.81	1.00	0.05	22 (0.41)	0.06	0.04	1.56
5	0.93	500 (3.58)	1.11	1.06	1.05	0.08	55 (1.01)	0.11	0.12	0.88
6	1.32	822 (5.89)	1.59	1.64	0.97	0.14	100 (1.84)	0.19	0.18	1.04
7	1.91	1292 (9.26)	2.39	2.43	0.98	0.25	236 (4.35)	0.38	0.36	1.06
8	2.98	2143 (15.36)	3.88	3.85	1.01	0.56	634 (11.69)	0.94	0.99	0.95
9	5.00	3490 (25.01)	6.69	6.54	1.02	1.50	1494 (27.54)	2.64	2.60	1.02
10	8.75	4705 (33.72)	12.85	12.96	0.99	4.24	2858 (52.69)	8.39	8.04	1.04
Men										
1	—	196 (4.34)	0.40	0.42	0.95		7 (0.4)	0.02	0.01	2.32
2	0.43	200 (4.43)	0.45	0.46	0.98	0.03	18 (1.04)	0.03	0.04	0.83
3	0.47	239 (5.29)	0.49	0.49	1.00	0.04	21 (1.21)	0.04	0.04	1.07
4	0.51	221 (4.89)	0.53	0.48	1.11	0.05	22 (1.27)	0.06	0.05	1.10
5	0.56	254 (5.62)	0.58	0.54	1.08	0.06	36 (2.07)	0.07	0.07	1.04
6	0.61	339 (7.5)	0.65	0.71	0.92	0.08	51 (2.93)	0.10	0.08	1.28
7	0.69	345 (7.63)	0.76	0.70	1.08	0.12	83 (4.78)	0.16	0.15	1.06
8	0.84	488 (10.8)	0.97	0.92	1.06	0.20	180 (10.36)	0.29	0.30	0.96
9	1.16	796 (17.61)	1.54	1.47	1.05	0.40	378 (21.75)	0.67	0.66	1.02
10	2.11	1441 (31.89)	4.19	4.20	1.00	1.10	942 (54.2)	3.20	2.68	1.19

*10 year predicted risk %.

alongside the FRAX algorithm in a UK primary care population. Given that FRAX was developed in multiple selected cohorts from across the world, the marginally poorer performance is not unexpected.

Our new algorithm does not require any laboratory testing or clinical measurements. All the variables used within our algorithm will either be known to an individual patient or are collected as part of routine clinical practice and recorded within an individual patient's primary healthcare record. It can be implemented within clinical computer systems used in primary care and used to stratify the practice population by risk on a continuing basis without the need for manual data entry. The QFractureScores could therefore act as a basis for a systematic population based programme to identify high risk patients for further assessment and support the implementation of evolving clinical guidelines in the UK.

These algorithms, like those that predict cardiovascular disease,^{10 36} rely on routinely collected data and have the advantage that they are well calibrated to the setting in which they can be used and have good levels of discrimination. Assuming the effectiveness and cost effectiveness of suitable interventions found in randomised controlled trials^{39 40} extend to unselected high risk patients from primary care, the QFractureScores could be used at a population level to identify patients at high risk of fracture who might benefit from more detailed assessment regarding

potential interventions to reduce their risk. At the level of the individual patient, the algorithm can be used for self assessment in a web based calculator (www.qfracture.org), which is similar to the website for self assessment of cardiovascular disease derived from the same database (www.qrisk.org). It can help inform patients regarding their absolute individual risk so that they can have better information on which to base treatment decisions. Some interventions can prevent osteoporotic fracture in high risk patients. Daily supplementation with vitamin D₃ and calcium reduces rates of hip fracture among high risk older patients in institutional care.⁴¹ Bisphosphonates reduce hip and other fracture rates in community dwelling older women aged under 80.⁴² Hip protectors seem to reduce the incidence of hip fractures in institutional care, provided that compliance and adherence are achieved.⁴³

Hormone replacement therapy and fracture risk

Our study also provides some information on risks associated with different types and doses of HRT. We have shown an overall protective effect of HRT with a decreased risk with unopposed oestrogen. The effect is more marked for vertebral, distal radial, and hip fractures combined rather than hip fracture alone, which is probably because of lower numbers of patients with hip fracture by individual type of HRT. Our findings are consistent with those from the Women's Health

Initiative study and other studies.⁴⁴⁻⁴⁶ The loss of the protective effect of HRT on risk of fracture after stopping treatment is consistent with some^{46 47} but not other studies.⁴⁸ Our study is larger than previous studies^{44 49} and has a longer follow-up.⁴⁴ It also includes a wider age range than in other studies, in which the analysis is restricted to women 10 years beyond the usual age for the menopause and those who have more risk factors and fewer menopausal symptoms.⁴⁴ Our study population is less likely to be biased than a clinical trial and so the results should generalise well to the general population of women in the UK deciding whether or not to start or continue HRT.

Validation

We validated the QFractureScore in an independent sample of general practices from which data had not been used to develop the algorithm. The QFractureScore has good discrimination (that is, ability to separate out those who did and did not subsequently develop a fracture) and explains over 60% of the variation for hip fracture. The D statistic, which is a measure of discrimination appropriate for survival type data, was substantially higher than in our cardiovascular disease algorithm⁵⁰ and than that reported in some other studies using the D statistic.³⁷ This increases the likelihood that the algorithm will accurately predict risk for an individual patient. This improved performance of the hip fracture algorithm compared with overall fracture and other outcomes is probably because of the

strong association between risk of hip fracture and age and might also reflect some stronger associations with other risk factors and a more accurate diagnosis, thus reducing misclassification and potential underestimation of associations.

A potential limitation of our validation might be a degree of overoptimism because, although we used a completely physically discrete set of general practices for the validation, these practices use the same clinical computer system (EMIS) as those used to derive the algorithm. The EMIS system, however, is currently in use in 60% of UK general practices and so the QFractureScore is at least likely to perform well for over half of the UK's population. A more stringent test of performance would involve practices using a different clinical computer system and such a study using the THIN database is currently under way. Validation of other disease algorithms in the THIN database have shown similar levels of performance to the validation undertaken in a one third sample of the QResearch database.^{16 36}

Comparison with other risk prediction algorithms

While there is consensus on the need to develop more accurate estimates based on absolute as well as relative risk, there is no widely used standard method for assessing risk of hip fracture among primary care populations in the UK.⁵¹ Unlike FRAX, the QFractureScore can also be used in younger patients (including those aged 30-39) and can be used to estimate risk at one,

Table 10 | Predicted and observed risks for hip fracture at 10 years in patients aged 40-85 by 10th of predicted risk using QFractureScore and FRAX scores

	QFractureScore			FRAX		
	Mean predicted risk (%)	Observed risk (%)	Ratio predicted/observed	Mean predicted risk (%)	Observed risk (%)	Ratio predicted/observed
Women						
1	0.05	0.02	2.47	0.16	0.08	2.03
2	0.08	0.10	0.81	0.16	0.08	2.03
3	0.12	0.14	0.86	0.30	0.17	1.76
4	0.18	0.14	1.30	0.40	0.25	1.60
5	0.29	0.32	0.90	0.54	0.33	1.65
6	0.51	0.47	1.08	0.83	0.61	1.36
7	0.97	1.03	0.95	1.37	1.06	1.29
8	2.01	1.98	1.01	2.46	1.99	1.24
9	4.14	4.30	0.96	4.74	4.34	1.09
10	9.87	9.40	1.05	10.07	9.33	1.08
Men						
1	0.04	0.04	1.04	0.10	0.06	1.66
2	0.06	0.06	1.02	0.10	0.06	1.66
3	0.08	0.08	1.01	0.20	0.11	1.82
4	0.11	0.07	1.53	0.20	0.11	1.82
5	0.14	0.15	0.96	0.30	0.17	1.76
6	0.21	0.19	1.09	0.40	0.24	1.67
7	0.32	0.34	0.94	0.59	0.34	1.72
8	0.56	0.46	1.21	0.98	0.52	1.88
9	1.16	1.38	0.84	1.76	1.36	1.30
10	4.12	3.39	1.21	3.87	3.31	1.17

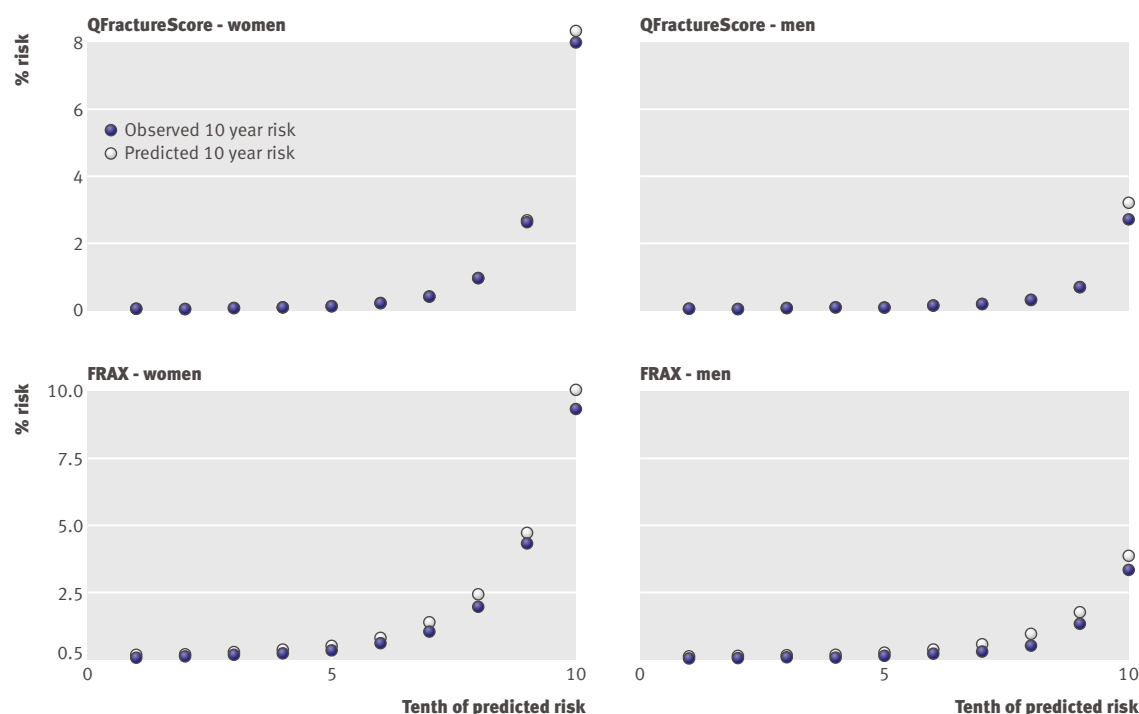


Fig 2 | Predicted and observed risk of hip fracture with QFractureScore and FRAX

two, five, and 10 years rather than just 10 years as with FRAX. Our new algorithm includes traditional variables included in the FRAX algorithm, such as age, smoking, alcohol, rheumatoid arthritis, parental history of hip fracture, and some secondary causes of osteoporosis. In addition, our algorithms include other risk factors as separate variables including type 2 diabetes, recorded history of falls, cardiovascular disease, liver disease, malabsorption, other endocrine disorders, use of tricyclic antidepressants, and type of HRT and menopausal symptoms in women. By including more detailed variables we hypothesise that the QFractureScores will be better at estimating risk for the individual patient by taking account of more information regarding the patient's history. Our algorithm differs from FRAX because the QFractureScores predict risk among patients without a recorded history of previous fracture, whereas FRAX can be used to predict future risk of fracture in those already known to have a previous fracture.

Our analysis shows some improved discrimination of the QFractureScore compared with the FRAX algorithm for hip fracture, based on the D statistic, which had values that were 0.11 higher in women and 0.17 higher in men. A difference in D statistic of 0.1 or more can indicate an important difference in prognostic separation of survival curves between two risk algorithms.³⁷ It is important to note, however, that the QFractureScore was developed in a primary care population and so would be expected to perform better in this setting.

One potential limitation of the QFractureScores compared with FRAX is that they don't include measurement of bone mineral density, whereas FRAX has two

versions, one with and one without this measurement. This potentially limits the value of the QFractureScores when bone mineral density is known but does mean that the score can be applied without the need for expensive and inconvenient tests to identify high risk patients who might then benefit from further investigation. Another potential limitation of the QFractureScores is that they are more complex than FRAX and might be more difficult to implement. The main use of the QFractureScores, however, is likely to be integrated into general practice clinical computer systems, as well as a web based calculator, where software can automatically extract the necessary variables, perform the calculations, and present the results to the clinician and individuals as appropriate. Also, a score that includes more variables is likely to better predict risk for an individual especially if that person has complex comorbidities. Future research should address whether a targeted approach to case finding based on the QFractureScores results in measurable clinical benefit.

Our algorithm also improves on the recent algorithm based on the Women's Health Initiative cohort²⁵ as it is estimated over a longer period than five years and includes additional variables, such as HRT, that are known to affect risk of fracture. It also has improved validation statistics; our ROC value for hip fracture in women was 0.890 (0.889 to 0.892), which is substantially higher than the value of 0.80 (0.77 to 0.82) reported in the Women's Health Initiative.²⁵

Methodological considerations

Generalisability and measurement of outcomes

A particular strength of our study is its prospective cohort design based on the analysis of a large

WHAT IS ALREADY KNOWN ON THIS TOPIC

Osteoporotic fracture is a major cause of morbidity, and interventions exist that can help reduce risk of fracture

Several international guidelines suggest a targeted approach for identifying high risk patients likely to benefit from interventions based on a 10 year absolute fracture risk

Risk prediction algorithms tend to perform best when they are developed in the clinical setting in which they will be applied

WHAT THIS STUDY ADDS

These new risk prediction algorithms (QFractureScores) for osteoporotic fracture and hip fracture do not require laboratory measurement and so can be used in primary care or for individual self assessment (www.qfracture.org)

The new algorithms include additional variables and were developed in and could be used in large representative primary care populations

The validation statistics, especially for the hip fracture algorithm, suggest that the QFractureScores are likely to be effective at identifying patients at high risk of fracture within primary care in the UK and showed improved performance compared with FRAX

representative population from a validated database. Our main outcome was hip, vertebral, or distal radial fracture recorded by a clinician on the clinical computer system. Similar studies using similar databases have confirmed the diagnosis of hip fracture on computerised general practice records in over 90% of cases.²⁶ Our rates of hip fracture were similar to those obtained in other similar general practice databases such as GPRD.⁵²⁻⁵⁴ Our rates tend to be higher than some population based cohorts,^{25 46} which might be because of some under-ascertainment due to self reported events reported by questionnaire rather than by analysis of data prospectively recorded on the patients' medical records, as in our study.

In our study, 38% of osteoporotic fractures were hip fractures, which is similar to the figure reported elsewhere.⁵⁴ There might be under-ascertainment or under-recording of vertebral fracture as it is not always associated with pain and loss of function so the patient might not present to the general practitioner. Failure to identify and record a diagnosis on the computer when a patient is identified is possible and is part of the justification for having a targeted approach.⁵⁵

Predictor variables

Our study has good validity as our hazard ratios for risk of hip fracture with use of corticosteroids, alcohol, smoking, diabetes, and presence of a parental history of osteoporosis were similar to those found in other studies.^{20 22 23 25 45 56 57} In particular, our analysis supports a dose-response relation for current smokers with lower risks among former smokers.¹⁹⁻²¹ We also found no association between hip fracture and deprivation, which confirmed findings reported elsewhere.⁵⁸

We have not included bone mineral density as the score is seldom recorded in general practice records, it is likely to be measured in only a selected high risk population, and it is costly to measure. The score could, however, be used to select high risk patients for measurement of bone mineral density as part of

their assessment after identification of their high risk status for fracture. Previous fracture was not included as these people have already experienced the outcome of interest, and it could be argued that they are automatically at high risk and all such patients should be managed as high risk patients in a secondary prevention context.

Sources of bias and unmeasured confounding

As with all epidemiological studies, we need to consider potential sources of bias and confounding. Our predictor variables were recorded by clinicians on the clinical computer system before the diagnosis of fracture and so will not have been subject to recall bias. Some of our predictor variables are objective clinical diagnoses (such as type 2 diabetes, cardiovascular disease, or rheumatoid arthritis), others are directly measured values (such as BMI). Several variables, however, are those reported by patients, such as alcohol, smoking, and parental history of hip fracture. As such these variables might be subject to information or reporting bias (patients, for example, might not accurately report their alcohol intake or use of cigarettes or might not be asked about or be aware of a relevant family history). As the QFractureScores are intended for use within general practice clinical computer systems, however, similar conditions will apply and so the variables incorporated in the algorithm have intrinsic face validity. We used the entire eligible population registered with a random two thirds sample of practices contributing to the QResearch database from England and Wales. Consequently, the population is unlikely to be affected by selection bias, in contrast with purpose designed clinical cohorts or clinical trials.^{46 59}

We did not have objective measurements of some other factors that might affect fracture risk, such as physical activity and ethnicity. The former is not reliably recorded on clinical computer systems, and analysis of the latter was limited in this study because of low numbers of elderly patients with fractures from different ethnic groups.

Missing data

Another potential limitation of our study is that some patients had missing values for alcohol use, BMI, or smoking status. We therefore used the technique of multiple imputation to substitute missing values rather than exclude these patients as this is a less biased approach that makes the most efficient use of available data.^{32 34} For other variables in the algorithm we assumed that if there was no recorded value of a diagnosis, prescription, or family history then the patient did not have that exposure, which might have led to some misclassification.

Conclusions

Our new risk prediction algorithms for osteoporotic fracture and hip fracture do not require laboratory measurement and can be readily used in primary care or for individual self assessment (www.qfracture.org). The algorithms potentially improve on other algorithms by

including additional variables not included in traditional scores. The validation statistics for the hip fracture algorithm suggest that the models are likely to be at least as effective at identifying patients at high risk of hip fracture within primary care as the FRAX algorithm. Further validation studies are needed to test the performance of these algorithms in independent populations that are representative of the setting where the algorithms are likely to be used.

We acknowledge the contribution of EMIS and EMIS practices contributing to the QResearch database.

Contributors: JHC designed the study, obtained approvals, reviewed the literature, prepared the data, developed and tested the algorithms, undertook the primary analysis and interpretation, and wrote the first and subsequent draft manuscript. CC contributed to the analysis, development and testing of the algorithms, and interpretation and drafting of the manuscript. JHC is the guarantor.

Funding: This study was funded by David Stables (medical director of EMIS) as part of a larger study examining risks and benefits of HRT.

Competing interests: JHC is codirector of QResearch, a not-for-profit organisation that is a joint partnership between the University of Nottingham and EMIS (leading supplier of IT for 60% of general practices in the UK). EMIS may implement the QFractureScore within its clinical system. JHC is also director of ClinRisk and CC is a consultant statistician for ClinRisk. ClinRisk produces software to ensure the reliable and updatable implementation of clinical risk algorithms within clinical computer systems to help improve patient care. This work and any views expressed within it are solely those of the co-authors and not of any affiliated bodies or organisations.

Ethical approval: The proposal was approved by the QResearch Scientific Board and is therefore approved by the Trent multicentre research ethics committee.

Data sharing: The Read codes used to define the outcomes and further information on multiple imputation are available from the authors. The algorithms have been published as open source software, which is available from www.qfracture.org.

- National Institute for Health and Clinical Excellence. Systematic review of clinical effectiveness prepared for the guideline: "osteoporosis assessment of fracture risk and the prevention of osteoporotic fractures in individuals at high risk." NICE, 2008.
- Kanis JA, Burlet N, Cooper C, Delmas PD, Reginster JY, Borgstrom F, et al. European guidance for the diagnosis and management of osteoporosis in postmenopausal women. *Osteoporos Int* 2008;19:399-428.
- Dawson-Hughes B, Tosteson A, Melton L, Baim S, Favus M, Khosla S, et al. Implications of absolute fracture risk assessment for osteoporosis practice guidelines in the USA. *Osteoporos Int* 2008;19:449-58.
- Kanis JA, McCloskey EV, Johansson H, Strom O, Borgstrom F, Oden A. Case finding for the management of osteoporosis with FRAX—assessment and intervention thresholds for the UK. *Osteoporos Int* 2008;19:1395-408.
- National Osteoporosis Guideline Group. Guideline for the diagnosis and management of osteoporosis in postmenopausal men and women from the age of 50 in the UK. National Osteoporosis Guideline Group, 2008.
- National Institute for Health and Clinical Excellence. Alendronate, etidronate, risedronate, raloxifene and strontium ranelate for the primary prevention of osteoporotic fragility fractures in postmenopausal women. NICE, 2008.
- Kanis JA, Johnell O, Oden A, Johansson H, McCloskey E. FRAX and the assessment of fracture probability in men and women from the UK. *Osteoporos Int* 2008;19:385-97.
- Wainwright SA, Marshall LM, Ensrud KE, Cauley JA, Black DM, Hillier TA, et al. Hip fracture in women without osteoporosis. *J Clin Endocrinol Metab* 2005;90:2787-93.
- Black DM, Steinbuch M, Palermo L, Dargent-Molina P, Lindsay R, Hoseney MS, et al. An assessment tool for predicting fracture risk in postmenopausal women. *Osteoporos Int* 2001;12:519-28.
- Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;336:1475-82.
- World Health Organization. WHO scientific group on the assessment of osteoporosis at primary health care level. WHO, 2004.
- Kanis J, Oden A, Johnell O. Acute and long-term increase in fracture risk after hospitalization for stroke. *Stroke* 2001;32:702-6.
- Michaelsson K, Baron JA, Farahmand BY, Johnell O, Magnusson C, Persson PG, et al. Hormone replacement therapy and risk of hip fracture: population based case-control study. The Swedish Hip Fracture Study Group. *BMJ* 1998;316:1858-63.
- Hubbard R, Farrington P, Smith C, Smeeth L, Tattersfield A. Exposure to tricyclic and selective serotonin reuptake inhibitor antidepressants and the risk of hip fracture. *Am J Epidemiol* 2003;158:77-84.
- Kush RD, Helton E, Rockhold FW, Hardison CD. Electronic health records, medical research, and the Tower of Babel. *N Engl J Med* 2008;358:1738-40.
- Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. *Heart* 2008;94:34-9.
- Willi C, Bodenmann P, Ghali WA, Faris PD, Cornuz J. Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA* 2007;298:2654-64.
- De Laet C, Kanis JA, Oden A, Johanson H, Johnell O, Delmas P, et al. Body mass index as a predictor of fracture risk: a meta-analysis. *Osteoporos Int* 2005;16:1330-8.
- Vestergaard P, Mosekilde L. Fracture risk associated with smoking: a meta-analysis. *J Intern Med* 2003;254:572-83.
- Cornuz J, Feskanich D, Willett WC, Colditz GA. Smoking, smoking cessation, and risk of hip fracture in women. *Am J Med* 1999;106:311-4.
- Hollenbach KA, Barrett-Connor E, Edelstein SL, Holbrook T. Cigarette smoking and bone mineral density in older men and women. *Am J Public Health* 1993;83:1265-70.
- Diaz MN, O'Neill TW, Silman AJ. The influence of family history of hip fracture on the risk of vertebral deformity in men and women: the European Vertebral Osteoporosis Study. *Bone* 1997;20:145-9.
- Kanis JA, Johansson H, Johnell O, Oden A, De Laet C, Eisman JA, et al. Alcohol intake as a risk factor for fracture. *Osteoporos Int* 2005;16:737-42.
- Spector TD, Hall GM, McCloskey EV, Kanis JA. Risk of vertebral fracture in women with rheumatoid arthritis. *BMJ* 1993;306:558.
- Robbins J, Aragaki AK, Kooperberg C, Watts N, Wactawski-Wende J, Jackson RD, et al. Factors associated with 5-year risk of hip fracture in postmenopausal women. *JAMA* 2007;298:2389-98.
- Van Staa T, Leufkens H, Abenhaim L, Zhang B, Copper C. Use of oral corticosteroids and risk of fractures. *J Bone Miner Res* 2000;15:993-1000.
- Weakliem DL. A critique of the Bayesian information criterion for model selection. *Sociol Methods Res* 1999;27:359-97.
- Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol* 1999;28:964-74.
- Schafer J, Graham J. Missing data: our view of the state of the art. *Psychol Methods* 2002;7:147-77.
- Group TAM. Academic medicine: problems and solutions. *BMJ* 1989;298:573-9.
- Steyerberg EW, van Veen M. Imputation is beneficial for handling missing data in predictive models. *J Epidemiol Community Health* 2007;60:979.
- Moons KGM, Donders RART, Stijnen T, Harrell FJ. Using the outcome for imputation of missing predictor values was preferred. *J Epidemiol Community Health* 2006;59:1092.
- Royston P. Multiple imputation of missing values: Update of ice. *Stata J* 2005;5:527-36.
- Royston P. Multiple imputation of missing values. *Stata J* 2004;4:227-41.
- Clark T, Altman D. Developing a prognostic model in the presence of missing data: an ovarian cancer case study. *J Epidemiol Community Health* 2003;56:28-37.
- Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ* 2007;335:136.
- Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. *Stat Med* 2004;23:723-48.
- Royston P. Explained variation for survival models. *Stata J* 2006;6:1-14.
- McCloskey EV, Beneton M, Charlesworth D, Kayan K, deTakats D, Dey A, et al. Clodronate reduces the incidence of fractures in community-dwelling elderly women unselected for osteoporosis: results of a double-blind, placebo-controlled randomized study. *J Bone Miner Res* 2007;22:135-41.
- Kanis JA, Borgstrom F, Johnell O, Oden A, Sykes D, Jonsson B. Cost-effectiveness of raloxifene in the UK: an economic evaluation based on the MORE study. *Osteoporos Int* 2005;16:15-25.

- 41 Chapuy M, Arlot M, Duboeuf F, Brun J, Crouzet B, Arnaud S, et al. Vitamin D3 and calcium to prevent hip fractures in elderly women. *NEJM* 1992;327:1637-42.
- 42 McClung MR, Geusens P, Miller PD, Zippel H, Bensen WG, Roux C, et al. Effect of risedronate on the risk of hip fracture in elderly women. Hip Intervention Program Study Group. *N Engl J Med* 2001;344:333-40.
- 43 Parker M, Gillespie LD, Gillespie WJ. Hip protectors for preventing hip fracture in the elderly (Cochrane review). Update Software, 2001.
- 44 Writing Group for the Women's Health Initiative Investigators. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results from the Women's Health Initiative randomized controlled trial. *JAMA* 2002;288:321-33.
- 45 Howard L, Kirkwood G, Leese M. Risk of hip fracture in patients with a history of schizophrenia. *Br J Psychiatry* 2007;190:129-34.
- 46 Banks E, Beral V, Reeves G, Balkwill A, Barnes I. Fracture incidence in relation to the pattern of use of hormone therapy in postmenopausal women. *JAMA* 2004;291:2212-20.
- 47 Heiss G, Wallace R, Anderson GL, Aragaki A, Beresford SAA, Brzyski R, et al. Health risks and benefits 3 years after stopping randomized treatment with estrogen and progestin. *JAMA* 2008;299:1036-45.
- 48 Bagger YZ, Tanko LB, Alexandersen P, Hansen HB, Mollgaard A, Ravn P, et al. Two to three years of hormone replacement treatment in healthy women have long-term preventive effects on bone mass and osteoporotic fractures: the PERF study. *Bone* 2004;34:728-35.
- 49 Hulley S, Furberg C, Barrett-Connor E, Cauley J, Grady D, Haskell W, et al. Noncardiovascular disease outcomes during 6.8 years of hormone therapy: heart and estrogen/progestin replacement study follow-up (HERS II). *JAMA* 2002;288:58-66.
- 50 Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation of QRISK2. *BMJ* 2008;336:1475-82.
- 51 Moayyeri A. Comment on Kanis et al: the use of clinical risk factors enhances the performance of BMD in the prediction of hip and osteoporotic fractures in men and women. *Osteoporos Int* 2008;19:591-2.
- 52 Card T, West J, Hubbard R, Logan RFA. Hip fractures in patients with inflammatory bowel disease and their relationship to corticosteroid use: a population based cohort study. *Gut* 2004;53:251-5.
- 53 Yang Y-X, Lewis JD, Epstein S, Metz DC. Long-term proton pump inhibitor therapy and risk of hip fracture. *JAMA* 2006;296:2947-53.
- 54 Van Staa T, Abenhaim L, Cooper C, Zhang B, Leufkens H. Public health impact of adverse bone effects of oral corticosteroids. *Br J Clin Pharmacol* 2001;51:601-7.
- 55 Holt T, Stables D, Hippisley-Cox J, O'Hanlon S, Majeed A. Identifying undiagnosed diabetes: cross-sectional survey of 3.6 million patients' electronic records. *Br J Gen Pract* 2008;58:192-6.
- 56 Van Staa TP, Leufkens HGM, Cooper C. Use of inhaled corticosteroids and risk of fractures. *J Bone Miner Res* 2001;16:581-8.
- 57 Kanis JA, Johnell O, Oden A, Johansson H, De Laet C, Eisman JA, et al. Smoking and fracture risk: a meta-analysis. *Osteoporos Int* 2005;16:155-62.
- 58 West J, Hippisley-Cox J, Coupland C, Price G, Groom L, Kendrick D, et al. The effect of deprivation, rurality and ethnicity on falls and hip fracture in elderly people in Trent 1992-1997: an ecological study. *Public Health* 2004;8:576-81.
- 59 Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med* 2000;342:1878-86.

Accepted: 8 July 2009