

**Predicting the risk of falls in patients with an indication for antihypertensive treatment: development and external validation of the STRATIFY-Falls prediction model**

Lucinda Archer (0000-0003-2504-2613),<sup>1</sup> *lecturer*, Constantinos Koshiaris, *statistician*,<sup>2</sup> Sarah Lay-Flurrie, *senior statistician*,<sup>2</sup> Kym IE Snell (0000-0001-9373-6591) *senior lecturer*,<sup>1</sup> Richard D Riley, *professor of biostatistics*,<sup>1</sup> Richard Stevens, *associate professor*,<sup>2</sup> Amitava Banerjee, *professor of clinical data science* (0000-0001-8741-3411)<sup>3</sup>, Juliet A Usher-Smith, *assistant professor* (0000-0002-8501-2531)<sup>4</sup>, Andrew Clegg, *professor of geriatric medicine*<sup>5</sup>, Rupert A Payne, *senior lecturer* (0000-0002-5842-4645)<sup>6</sup>, FD Richard Hobbs, *Nuffield professor of primary care*<sup>2</sup>, Richard J McManus, *professor of primary care research* (0000-0003-3638-028X)<sup>2</sup>, James P Sheppard (0000-0002-4461-8756), *associate professor*,<sup>2</sup> on behalf of the STRATIFY investigators\*

<sup>1</sup> Centre for Prognosis Research, School of Medicine, Keele University

<sup>2</sup> Nuffield Department of Primary Care Health Sciences, University of Oxford

<sup>3</sup> Institute of Health Informatics, University College London, London, UK

<sup>4</sup> Primary Care Unit, Department of Public Health and Primary Care, University of Cambridge, UK

<sup>5</sup> Academic Unit for Ageing and Stroke Research, Bradford Institute for Health Research, University of Leeds, UK

<sup>6</sup> Centre for Academic Primary Care, Population Health Sciences, University of Bristol, Bristol, UK

\*The STRATifying Treatments In the multi-morbid Frail elderly (STRATIFY) investigators include the authors and the following:

John Gladman, professor of medicine of older people, School of Medicine, University of Nottingham; Simon Griffin, professor of primary care, Department of Public Health and

Primary Care, Primary Care Unit, University of Cambridge; and Margaret Ogden, Patient and Public Involvement advisor.

**Corresponding author:** James P Sheppard (0000-0002-4461-8756)

**Email:** james.sheppard@phc.ox.ac.uk

**Telephone:** +44 1865 617192

**Address:** Nuffield Department of Primary Care Health Sciences, Radcliffe Primary Care Building, Radcliffe Observatory Quarter, University of Oxford, Oxford, OX2 6GG, UK

**Word count:** 5,203

**Number of references:** 56

**Number of tables:** 3

**Number of figures:** 5

**Number of boxes:** 1

## **Study summary**

### **What is already known on this topic**

- Serious falls are possible side effect of antihypertensive treatment, which can adversely affect patients' quality of life and increase the risk of hospitalisation, especially in older people with frailty.
- Existing tools that estimate an individual's risk of falls have been shown to be at high risk of bias with only moderate discriminative ability.

### **What this study adds**

- In the present study, a clinical prediction model for the risk of falls up to 10 years in the future was developed and externally validated, incorporating commonly recorded patient characteristics, co-morbidities and medications, in patients with an indication for antihypertensive treatment.
- Upon external validation, the model discriminated well between patients who went on to have a serious fall, but calibration indicated under-prediction of risk.
- Nevertheless, a decision curve analyses suggests the model has clinical utility, and so may be useful to identify patients whose risk of a fall is high, who may require closer monitoring or may benefit from closer monitoring or early intervention to prevent future falls.

## **Abstract**

**Objectives:** To develop and externally validate a clinical prediction model to identify the risk of hospitalisation or death from a fall in patients eligible for antihypertensive treatment.

**Design:** Retrospective cohort study using primary care data from electronic health records contained within the Clinical Practice Research Datalink (CPRD).

**Participants:** Patients aged 40 years or older, with at least one blood pressure measurement between 130-179 mm Hg.

**Main outcome measure:** First serious fall, defined as hospitalisation or death with a primary diagnosis of a fall within 10 years of the index date.

**Analysis:** Model development was conducted using a Fine-Gray approach in data from CPRD GOLD, accounting for the competing risk of death from other causes, with subsequent recalibration at 1, 5 and 10 years using pseudo-values. External validation was conducted using data from CPRD Aurum, with performance assessed through calibration curves and the Observed/Expected (O/E) ratio, C-statistic, and D-statistic, pooled across GP practices, and clinical utility using decision curve analysis at thresholds around 10%.

**Results:** Analysis included 1,773,224 patients (62,691 serious falls) from CPRD GOLD used in model development, and 3,805,322 (206,956 serious falls) from CPRD Aurum in the external validation. The final model consisted of 24 predictors, including age, sex, ethnicity, alcohol consumption, living in an area of high social deprivation, a history of falls, multiple sclerosis and prescriptions of antihypertensives, antidepressants or hypnotics/benzodiazepines. Upon external validation, the recalibrated model showed good discrimination, with pooled C-statistics of 0.833 (95%CI: 0.832 to 0.835) and 0.833 (95%CI: 0.831 to 0.835) at 5 and 10 years respectively. Original model calibration was poor on visual inspection and whilst this was improved with recalibration, under-prediction of risk remained (O/E at 10 years 1.839, 95%CI: 1.811 to 1.865). Nevertheless, decision curve analysis suggests potential clinical utility, with net benefit larger than other strategies.

**Conclusions:** This prediction model uses commonly recorded clinical characteristics and distinguishes well between patients at high and low risk of falls in the next 1-10 years.

Although miscalibration was evident on external validation, the model still had potential clinical utility around risk thresholds of 10%, and so could be useful in routine clinical practice to help identify those at high risk of falls, who might benefit from closer monitoring or early intervention to prevent future falls. Further studies are needed to explore the appropriate thresholds which maximise the model's clinical utility and cost-effectiveness, and to examine whether recalibration is possible in local settings.

**Word Count:** 400 words

## **Print abstract**

**Study question:** This study aimed develop and externally validate a clinical prediction model to identify the risk of hospitalisation or death from a fall, in patients indicated for antihypertensive treatment.

**Methods:** Retrospective cohort designs were used for model development and external validations, using two separate sets of UK primary care data from electronic health records contained within the Clinical Practice Research Datalink (CPRD). Model development was conducted in data from CPRD GOLD, using a competing risk model to account for the risk of death from other causes. External validation was conducted using data from CPRD Aurum.

**Study answer and limitations:** The model incorporates routinely recorded information including a history of previous falls, multiple sclerosis, heavy alcohol consumption, high deprivation score and prescribed medications, which were all strong predictors of subsequent serious falls, conditional on the other model variables. Upon external validation, the model discriminated well between patients who went on to have a serious fall and those who did not (C-statistic at 10 years 0.83, 95%CI 0.83 to 0.84), but calibration indicated under-prediction of risk, particularly in those at higher risk of serious falls (O/E ratio at 10 years 1.84, 95%CI: 1.81 to 1.87).

**What this study adds:** Despite mis-calibration, analyses suggest the model has clinical utility, and so may be useful to identify patients whose risk of a fall is high, who may benefit from closer monitoring or early intervention such as deprescribing to prevent future falls.

**Funding, competing interests, and data sharing:** Authors had financial support from the Wellcome Trust, Royal Society, and National Institute for Health Research for the submitted work. Data were obtained via a CPRD institutional licence. Requests for data sharing should be made directly to the CPRD. The algorithm is freely available for research use and can be downloaded from <https://process.innovation.ox.ac.uk/software/>.

## Introduction

The proportion of older adults in the population is rising (1), and as people age, they become at increased risk of falls (2, 3) which can result in serious injury and long-term disability (4).

In England, falls are associated with approximately 235,000 emergency hospital admissions in the over 65s and cost the National Health Service more than £2.3 billion every year (5-7).

There are many risk factors for falls, primarily related to individuals who develop co-morbidities and frailty (2, 3, 8-10). A key modifiable risk factor is prescribed medications, including those that lower blood pressure (11-13). These are prescribed to reduce the risk of cardiovascular disease (CVD), and whilst they are very effective, typically many patients require treatment over a number of years to prevent a smaller number of events (14). Data from randomised controlled trials shows that antihypertensives are associated with an increased risk of hypotension and syncope, which may lead to falls (15). Indeed, observational studies examining patients with frailty and multi-morbidity suggest a direct association between antihypertensive treatment and falls (11, 16, 17).

In patients who are prescribed antihypertensives or other medications which significantly increase their falls risk, physicians may wish to consider altering or withdrawing treatment (i.e. deprescribing) (18), along with other interventions to reduce falls risk (e.g. advice regarding reduced alcohol consumption, falls prevention clinics, exercises)(7). However, identifying patients at high risk of falls is challenging. A 2021 systematic review of falls prediction models for use in the community identified a total of 72 models (10). The majority of these studies were deemed at high risk of bias and only three of the models were externally validated. These three validated models showed moderate discriminative ability with an area under the curve (AUC) of between 0.62 to 0.69. Calibration was only reported in seven studies based on internal validation, and was typically moderate to poor (10). A further primary analysis aiming to predict falls in a general practice population showed good

apparent discrimination for their model (with an AUC of 0.87), but calibration performance was not assessed and no external validation was performed (19).

To inform clinical decision making in primary care, both patients and physicians require better prediction models to accurately identify those at high risk of serious falls (defined as any fall resulting in hospitalisation or death), from the population of older adults who might be considered for antihypertensive therapy. This population includes patients with a recent high blood pressure reading, including those with a new diagnosis of hypertension, as well as those in whom intensification of treatment is being considered. The present study therefore aimed to use routinely collected data from electronic health records (EHRs) to develop and externally validate a clinical prediction model to estimate such an individual's risk of experiencing a fall resulting in hospitalisation or death within 10 years. This study is part of a broader research programme investigating the association between blood pressure lowering drugs and side effects: "STRATifying Treatments In the multi-morbid Frail elderly (STRATIFY): Antihypertensives".



## Methods

A retrospective observational cohort study was used to develop a prediction model for serious falls, using data from CPRD GOLD, which contains data from GP practices using Vision electronic health record software (Cegedim Healthcare Solutions, London, England). The model was externally validated using a second retrospective observational cohort comprising data from CPRD Aurum, containing data from GP practices using recording software from Egton Medical Information Systems (EMIS, Leeds, England). These data were linked to Office for National Statistics (ONS) mortality data, Hospital Episode Statistics (HES) and Index of Multiple Deprivation data (IMD). The protocol for this study was approved by the CPRD Independent Scientific Advisory Committee (ISAC) (protocol number 19\_042, see Appendix 5 in the supplementary material).

### Population

Patients were eligible if they were registered at a linked general practice in England, contributing to CPRD between 01/01/1998 and 31/12/2018. At the time of analysis, CPRD GOLD (development cohort) contained 4.4 million active patients from 674 GP practices, while CPRD Aurum (validation cohort) contained 7 million active patients from 738 practices. Both datasets have previously been shown to be representative of the patient population in England in terms of age, ethnicity, and deprivation status (20, 21). To avoid duplicate patients, where practices had switched from one recording system to the other during the study timeframe, we excluded all practices from CPRD Aurum (validation cohort) that were also present in CPRD GOLD (development) dataset (see figure 1).

Patients were considered eligible if they were aged 40 years or older (no upper age limit applied), registered to a CPRD “up-to-standard” practice (CPRD GOLD only) and had records available during the study period. Patients entered the cohorts at the time at which they became potentially eligible for antihypertensive treatment; i.e. at the time of their first systolic blood pressure reading  $\geq 130$  mmHg, after the study start date, and were followed for

up to 10 years. This blood pressure threshold was chosen to account for varying treatment initiation thresholds specified in different international hypertension guidelines (6). Patients with any systolic blood pressure reading greater than 180 mmHg were excluded from the cohort, as antihypertensive treatment would be indicated for these patients regardless of the risk of adverse events, unless clearly contraindicated for other reasons. All patient characteristics and model predictors were determined at the index date, defined as 12 months after cohort entry. The same eligibility criteria and characteristic determination methods were applied to both the development and validation cohorts.

### Outcomes

The primary outcome was defined as any in hospitalisation or death associated with a primary diagnosis of a fall within 10 years of the index date, the same time horizon as used for cardiovascular prediction models (22). Falls were based on International Classification of Diseases, tenth revision, (ICD10) codes documented in HES and ONS mortality data (applicable ICD10 codes shown in supplementary table S4.1). Pre-specified secondary outcomes were falls (defined in the same way) within 1 and 5 years of index date. This outcome definition was consistent across both the development and validation cohorts.

### Model predictors

Clinically relevant predictors of falls were identified from the literature and through expert clinical opinion (2, 7-9, 23). These included 30 predictors (44 predictor parameters), covering patient demographics (age, sex, ethnicity, area-based socioeconomic deprivation [IMD], systolic/diastolic blood pressure [SBP/DBP], body mass index [BMI]), clinical characteristics (total cholesterol, smoking, alcohol intake), comorbidities (previous falls, memory issues, mobility issues, history of stroke, multiple sclerosis, activity limitation, syncope, cataract) and prescribed medications (antihypertensives, opioids, hypnotics/benzodiazepines, antidepressants, anticholinergic medications) (see table S4.2 in the supplementary material). A recent literature review of falls clinical prediction tools by the National Institute for Health

and Care Excellence identified the need for frailty to be considered as a predictor in models for use in the community (24). We therefore also calculated a validated Electronic Frailty Index using the 36 comorbidities and conditions specified, including this index as a single covariate (25). Covariates were defined by any occurrence of relevant Read codes/SNOMED codes at any time point before the index date, with the exception of antihypertensives, which were defined as any prescription in the 12 months prior to the index date.

To ensure consistency with commonly used risk calculators (26, 27), our prediction models do not account for changes in prescriptions of medication type or amount over time, and as such give an estimation of falls risk assuming treatment assignment policy in any application setting is similar to that in the development data (28).

### Sample size

The pre-specified sample size calculation for model development was 2,194 participants (15,358 person-years), assuming a maximum of 40 predictors would be included in the final model (see extended methods in the supplementary material) (29). For the external validation, the estimated sample size required was 12,000 patients (with approximately 708 experiencing falls), sufficient to target a 95% confidence interval (CI) of width 0.2 around the estimate of the calibration slope (see extended methods)(30). The actual sample sizes in both development and validation cohorts far exceeded these estimates.

### Statistical analysis

Descriptive statistics were calculated for baseline characteristics in the model development and external validation cohorts separately.

### *Missing data*

Multiple imputation with chained equations was used to impute missing data in both the development and validation cohorts, with ten imputations generated for the development and validation datasets. Two separate and independent imputation procedures were used, one for model development and one for model validation. The imputation models included all model covariates within each dataset, along with the Nelson-Aalen estimator for the cumulative baseline cause-specific hazards for falls and for the competing event of death, and binary event indicators for each of these possible event types (31, 32). Where information was missing on the diagnosis of comorbidities or prescribed medications, it was assumed that no diagnosis or prescription was present. Predictor variables requiring imputation were cholesterol, ethnicity, deprivation score (validation cohort only), smoking status, and alcohol consumption.

Imputations were assessed for consistency by comparing density plots, histograms, and summary statistics across imputations and back to the complete values. The model coefficients and predictive performance measures were then estimated in each imputed dataset separately, before being combined across imputations using Rubin's Rules (33).

### *Model development*

Model development and internal validation was conducted by researchers at the University of Oxford (CK, JPS). Multivariable prediction models were fitted in each imputed dataset using a Fine-Gray sub-distribution hazard model, taking into account the competing risk of death by other causes (34). Accounting for the competing risk in this way aimed to avoid overestimation of the predicted probabilities of falls (35). Predictor effects in the model are reported as sub-distribution hazard ratios (SHR) with 95% confidence intervals, and the post-estimation baseline cumulative incidence for falls was estimated using a Breslow type estimator (34). Analyses were undertaken using the *fastcmprsk* package in RStudio (36). Automated variable selection methods were not used, since the variables were all predetermined based on the literature and expert opinion, and given the large sample size

would result in nearly all predictors having a statistically significant association with the outcome, regardless of effect size. To ensure a parsimonious model, variables with little or no association in multivariable analysis were excluded before fitting the final model.

Fractional polynomial terms were examined to identify the best fitting functional form of all continuous variables (37). Fractional polynomials (FPs) were identified separately within each imputed dataset and the most consistent transformation across the imputations was selected, choosing lower order FP terms wherever possible for the sake of parsimony. The selected FP format for each continuous variable was then forced into the model for all imputations to ensure consistency in coefficient estimation.

Interactions between age, sex and antihypertensives treatments were considered, but were excluded from the model development due to issues with stability or convergence, or for the sake of parsimony.

The proportional hazards assumption was checked for each predictor by examination of the Schoenfeld Residuals (38).

#### *Apparent validation using development data*

Observed outcome probabilities were defined using pseudo-values: jack-knife estimators representing an individual's contribution to the cumulative incidence function for falls.

Pseudo-values were generated separately in 50 groups by linear predictor value, for stability, and to account for the competing risk of death and non-informative right censoring (39, 40).

The model's apparent calibration performance was assessed using calibration plots comparing the observed to predicted risks at 1, 5 and 10 years. The calibration plots were produced using observed pseudo-values and included a smooth (non-linear) calibration curve to show apparent calibration across the spectrum of predicted risks (41), with 95%

confidence intervals. Plots were generated in each imputed dataset separately and were checked for consistency across imputations. A single, representative example is reported.

Where plots showed miscalibration, the original Fine-Gray model was recalibrated separately at each time point by fitting a generalised linear equation with a logit link function directly to the observed pseudo-values in the development dataset. The linear predictor from the original model was the only variable included in the recalibration model, which allowed for a non-linear recalibration effect using fractional polynomials.

### *External validation*

The external validation of the prediction model was conducted by researchers at Keele University (LA, KIES, RDR), independent of the model development team. The prediction model algorithms presented in Box 1 (both the original and final) were applied to each individual in the external validation cohort to give the predicted probabilities of experiencing a fall within 1, 5 and 10 years, taking account of the competing risk of death by other causes (42). Model calibration was assessed through comparison of predicted probabilities to observed pseudo-values, calculated in the external validation cohort in the same way as described above.

Predictive performance was quantified by calculating the Observed to Expected ratio (O/E), Harrell's C-statistic, Royston's D-statistic with its associated  $R^2$  statistic (43) and using calibration plots and curves. Calibration plots were generated separately in each imputed dataset and checked for consistency (one illustrative example is shown for each model). All measures were calculated in each imputed dataset separately and, where appropriate, combined across imputations using Rubin's Rules. Where Rubin's Rules did not apply (e.g. when the posterior distribution was not expected to be normal), performance was summarised across imputations using the median and IQR (44).

Heterogeneity in model performance across different GP practices was assessed using a random effects meta-analysis, using restricted maximum likelihood estimation (REML), given that the case mix and incidence of falls were expected to vary between practices (see extended methods in the supplementary material) (45). The O/E was pooled across practices on the natural log scale, the C-statistic on the logit scale (with the standard errors of logit-C calculated using the delta method), and the D-statistic on its original scale (46, 47). Pooled estimates are reported with prediction intervals (PI) to give an indication of expected model performance in a new GP practice.

Clinical utility was assessed by plotting the 1-, 5- and 10-year risk of falls against the 10-year risk of CVD, calculated using the Qrisk2 algorithm (22). Clinical utility was also examined using Net Benefit analysis, where the harms and benefits of using a model to guide treatment decisions were offset to assess the overall consequences of using the STRATIFY-Falls prediction models for clinical decision making (48). The original and final models were compared to one another at 5 and 10 years and to 'model-blind methods' of (a) introducing falls prevention measures (which may include deprescribing) for all patients or (b) not introducing falls prevention measures (continuing all treatment) for all patients, regardless of falls risk. Net benefit was assessed across the full range of possible threshold probabilities, with a falls risk above 10% at 10 years specified a priori as being a threshold of clinical interest, to align with current thresholds for an individual's CVD risk (22).

The same external validation methods as described above were employed in sub-groups by age (<65 years, ≥65 years), sex (female, male), and ethnicity (white, black, South Asian, other), to assess the models' predictive performance in these clinically relevant groups.

### Patient and Public Involvement

This study was developed and conducted with the help of our patient and public advisor Margaret Ogden. As a member of our study advisory group, they commented on the study

protocol and have been present in all team meetings discussing results and reporting. We also held a focus group with several older adults during the study to discuss broader issues related to drugs for cardiovascular disease prevention and adverse events, which informed the interpretation of this work.



## Results

### Study population characteristics

The flow of study participants for both the development and validation cohorts is shown in Figure 1. A total of 1,772,600 patients were included in the model development cohort (CPRD GOLD) with a mean age of 59 years (SD 13.2) and a mean systolic blood pressure at study inclusion of 144 mmHg (SD 12 mmHg) (table 1). The 10-year prevalence of falls was 3.5% (n= 62,691) with 10.3% of patients experiencing death by other causes before any fall could take place (n=181,731), and a median follow up time of 6.2 years (IQR: 2.6 to 10 years) across the cohort.

In total, 3,805,366 patients were included in the validation cohort, with 206,956 (5.4%) experiencing falls events during 10-year follow-up. A further 334,552 (8.8%) patients died during follow-up from unrelated causes, before any fall occurred. Median follow-up time in the validation cohort was 6.7 years (IQR: 2.7 to 10 years). Total cholesterol was missing in 48% of participants whilst ethnicity data were more complete in the validation cohort compared to the development cohort (81% vs 44% complete data).

### Model development

The original model consisted of 24 predictors, after the exclusion of variables with little or no association in multivariable analysis (table 2). Compared with men, women were more likely to experience a fall during follow-up (SHR 1.25, 95% CI: 1.23 to 1.27). Increasing age, white ethnicity, being a smoker, a heavy drinker, or being more deprived were associated with an increased risk of falls (table 2). Increasing frailty was one of the strongest predictors of falls, with an increased falls risk of 22% for approximately every four deficits accrued (SHR: 1.22, 95% CI: 1.20 to 1.23). Of the previous medical conditions examined, having a history of falls (SHR 1.32, 95% CI: 1.29 to 1.35) and multiple sclerosis were the strongest predictors of falls (SHR 1.71, 95% CI: 1.51 to 1.94). Angiotensin II receptor blockers (SHR 1.19 95% CI: 1.15 to 1.23), antidepressants (SHR 1.16 95% CI: 1.13 to 1.18), hypnotics/benzodiazepines (SHR

1.15 95% CI: 1.13 to 1.18), ACE inhibitors (SHR 1.12 95% CI: 1.10 to 1.14) and opioids (SHR 1.11 95% CI: 1.08 to 1.13) were the medications most strongly associated with falls. To ensure a parsimonious final model, SBP, DBP, BMI, activity limitation, syncope and cataract were excluded from the model due to a lack of effect on falls risk. No violations of the proportional hazards assumption were detected.

#### Internal validation and recalibration using pseudo values

At 5 and 10 years, apparent calibration plots in the model development data showed severe miscalibration, with under-prediction for those with a low predicted risk and substantial over-prediction for those with a high predicted risk (see supplementary figure S3.1). The original model was therefore recalibrated to the observed pseudo-values and this improved apparent calibration (in the model development data) considerably (Figure 2). Apparent calibration of the original model at 1 year was good, therefore recalibration was not required.

#### External validation

##### *Predictive performance*

Upon external validation, the original model showed excellent discrimination (table 3) but poor calibration (supplementary figure S3.1), with considerable heterogeneity across GP practices (supplementary figure S3.2). Recalibration of the model corrected miscalibration in the model development cohort, but under-prediction of risk was still present in the validation cohort (Figure 2). This miscalibration was less extreme than that of the original model, in the narrower range of predicted probabilities between 0 to 0.2. On average, the recalibrated model showed a pooled O/E ratio at 10 years of 1.839 (95% CI: 1.811 to 1.865, 95% PI: 1.284 to 2.638), suggesting that the observed incidence of falls would be around 84% (relatively) higher than expected when using the model to generate predictions. Under-prediction of 10-year falls risk was consistent across all subgroups, with the exception of the “other” ethnicity group, where both falls incidence and the Observed/Expected ratio were

considerably lower than in the full validation data (see extended results in supplementary material section 2.2).

Recalibration altered the ordering of participants' predicted probabilities only very slightly, thus discriminative ability of the recalibrated models remained excellent at each of the analysis timepoints, with C-statistics of 0.843 (95% CI: 0.841 to 0.844, 95% PI: 0.789 to 0.881 [5 years]) and 0.833 (95% CI: 0.831 to 0.835, 95% PI: 0.789 to 0.870 [10 years]) and D-statistic values of 1.894 (95% CI: 1.746 to 2.042, 95% PI: 1.75 to 2.04 [5 years]) and (1.597, 95% CI 1.472 to 1.721; 95% PI: 1.47 to 1.72 [10 years]; table 3). Model performance varied more among smaller practices, with more consistent performance seen as practice size increased (Figure 3).

The model's discriminative ability at 10 years was consistent across age and sex sub-groups (supplementary tables S2.1 and S2.2). The pooled C-statistic was lowest in those with white ethnicity at 0.796 (95% CI: 0.793 to 0.798) and highest among those with other ethnicity at 0.834 (95% CI: 0.830 to 0.839) (supplementary table S2.3).

### *Clinical utility analysis*

Net benefit and decision curve analysis of the original and recalibrated models indicated potential clinical utility at 5 and 10 years around the pre-defined threshold of 10% (Figure 4). At 10 years, basing clinical management decisions on predicted falls probabilities yielded a benefit over the “introduce falls prevention measures (which may include deprescribing) for all” and “not introducing falls prevention measures (continuing all treatment) for all patients” strategies when using a treatment decision threshold of 7% or higher from the original model, or a treatment decision threshold of 6% or higher from the final recalibrated model. Thus, for either model, when using our pre-specified treatment decision cut-off of 10% risk of falls at 10 years, we would expect a benefit to patients over-and-above model-blind

treatment strategies (usual care). This treatment decision threshold of 10% showed a net benefit in all sub-groups except other ethnicity, where a cut-off of at most 3% was required for the model to be superior to usual case for all (supplementary figure S2.6). In the analysis at 5 years, using a treatment decision threshold of 3% risk or higher gave a net benefit above continuing treatment for all, for both models.

In analyses comparing the risk of falls to the risk of CVD in CPRD GOLD, 198,654 (11%) patients had a high risk of falls (above 10%) but low risk of CVD (below 10%) over ten years (Figure 5). A further 128,458 (7%) patients were classified as high risk of both and 571,274 (32%) had a low falls risk but high risk of CVD.

## **Discussion**

### Summary of main findings

In the present study, a clinical prediction model was developed and externally validated to determine an individual's risk of experiencing a fall resulting in hospitalisation or death, within 10 years of being indicated for antihypertensive medication (due to raised blood pressure readings). The model incorporates routinely recorded information including a history of previous falls, multiple sclerosis, heavy alcohol consumption, high deprivation score and prescribed medication, which were all strong predictors of subsequent falls, conditional on the other model variables.

The final recalibrated model showed good discrimination upon external validation, suggesting that it can help distinguish those at a higher risk of falling, which may improve how physicians identify patients who might benefit from targeted fall prevention strategies, including multifactorial or exercise based interventions (49), and medication reviews. Calibration performance of the prediction model was inconsistent across the development and validation datasets, with miscalibration leading to under-prediction of fall risk across the full range of predicted probabilities. Nevertheless, such under-prediction of risk may be deemed acceptable if the model is intended to inform whether treatment should be stopped to avoid adverse effects, particularly if the treatment in question also carries benefits. Indeed, the clinical utility analysis showed that at risk thresholds around 10%, the net benefit of the model is higher than other strategies.

### Strengths and limitations of this work

Strengths of this work include the large, population-based cohorts used, incorporating routinely collected patient data which have been shown to be representative of the patients across England, suggesting that the findings could be generalised across this (or a similar)

population (20, 21). Analyses accounted for the competing risk of death in both model development and external validation, ensuring that falls risk was not over-estimated. This is particularly important in individuals with frailty and multiple long-term conditions, where an over-estimation of falls risk might preclude prescription of antihypertensive medication in those who could still derive benefit from continued treatment. The analysis method is superior to most prediction models in widespread use, which do not allow for competing risks or allow for competing risks with a cause-specific hazard framework (22). In these models, the stated risk of an event (cardiovascular disease, for example) is by design too high, as the actual risk of an event would be diminished by death from other (e.g., non-cardiovascular) causes, particularly in older people.

All data were derived from routine electronic health records, including the outcome definition of falls. Such a definition may not capture all events which might be included in the ProFaNE consensus definition of a fall (e.g. an unexpected event in which the participants come to rest on the ground, floor, or lower level) (50), and therefore the model results should be interpreted in this context. It is possible that some of these fall events were not reported or captured correctly within the electronic health record, therefore potentially underestimating the incidence of falls, which could have affected the performance of the model.

Assessments of the models' predictive performance were conducted across a range of general practices, with different case-mix and outcome prevalence, giving an indication of the expected spread of performance across a range of sub-populations. Model performance varied more among smaller practices, with more consistent performance seen as practice size increased. This reflects the increased uncertainty in the estimation of the predictive performance measures in practices of low sample sizes, many of which individually would have failed to meet the required sample size for this external validation. Prediction intervals

from meta-analyses across GP practices give an indication of how well our falls models would be expected to perform in new practices, helping to inform decisions on implementation in practice. In the present study, the prediction intervals were relatively narrow across a range of performance statistics, suggesting that the models would perform similarly in a new practice from a similar population.

All variables included in our model were predetermined based on the literature, although we did choose to exclude some variables at the model development stage that had exhibited a negligible effect on the outcome. These were excluded because they did not contribute substantially to model predictions and served to unnecessarily increase the complexity of the equation. We did not use statistical selection methods such as backwards or forwards elimination, as such can lead to overfitting. While our approach may have meant that some statistically significant (but clinically insignificant) predictors were excluded from the final model, these exclusions are unlikely to have led to overfitting given the large sample size, or been the reason for miscalibration in the external validation.

For these models, we defined binary variables for antihypertensive medication as any prescription within the year prior to (and including) the index date, without accounting for any changes to medication during follow-up. Not allowing for the time-varying nature of treatment could potentially affect the observed associations with falls risk, and so too the predicted risks obtained from the model. However, our model is intended to give a prediction of fall risk over the next 1-10 years, from a particular moment in time, in the context of current care. The latter is important, as for example, if a patient has low risk, then it means that current care (i.e., treatments and monitoring strategies over the next 1-10 years) is likely to be adequate for this individual. In contrast, if an individual's risk is high, it means that

current care is likely insufficient and additional or alternative approaches are potentially needed.

Calibration performance of the prediction model was inconsistent across the model development and validation datasets. Such miscalibration was surprising, as populations were similar across both datasets in terms of predictor distributions and the incidence of falls and of death (with the exception of self-reported characteristics such as smoking status, alcohol consumption and ethnicity, which may reflect difference in how these data are captured within the electronic health record systems which underlie these databases). Distributions of the linear predictor were also consistent across the development and validation datasets, suggesting miscalibration may be due to differences in the outcomes or the outcome recording or coding. This is representative of real life, where outcome definitions vary, and both models still exhibited useful discrimination and potential clinical utility across the full population for a range of treatment decision threshold probabilities, although the predicted risk for an individual may be different (miscalibrated) from their actual risk. Indeed, miscalibration was most evident in the 5-10% of patients with the highest predicted risk (those above a threshold of 10%), and for these patients physicians may interpret the exact predicted risks with caution, even though they can still be considered at higher risk.

#### Comparison with previous literature

A number of prediction models now exist to estimate an individual's risk of falls, including those for use in the community. A recent systematic review of development and validation studies identified a total of 69 existing models (10). These were typically poorly reported, with only 40 studies (56%) reporting discrimination statistics and only 7 studies (10%) reporting calibration. Only three models were externally validated. Discrimination was



reported with AUCs of between 0.49 to 0.87 for internally validated models and 0.62 to 0.69 for externally validated models. Calibration was moderately good, but presented in deciles of risk across a small range of risk thresholds (e.g. 0-10% (51)) making it difficult to determine how calibration varied across the full range of predicted probabilities. All studies were deemed at high risk of bias due to methods of analysis and outcome assessment along with restrictive eligibility criteria.

In contrast, our final model, reported in line with TRIPOD guidelines for reporting of clinical prediction models (52) (see supplementary table S4.3), showed excellent discrimination upon external validation with an AUC of 0.84. It demonstrated reasonable calibration across the low range of predicted risks typically examined by previous risk models (e.g., 0-10%), although miscalibration was present at higher predicted probabilities, there was still clinical utility based on the decision curve analysis. This suggests that the present model appears to be the most promising clinical prediction models for falls available to date, and may be effective in identifying individuals at high risk of falls from those in primary care with high blood pressure.

#### Implications for policy and practice

As patients age, their risk of a fall resulting in serious injury and long-term disability increases (4). Identifying those most at risk is therefore important to enable targeting of fall prevention strategies (7). The present model provides primary care physicians with a method of estimating fall risk using data routinely available in electronic health records and may have uses beyond predicting falls in patients being considered for antihypertensive treatment (53).

Among patients aged 40 years and older, with an indication for antihypertensive therapy due to raised blood pressure, the model was shown to distinguish well those at high risk of falls

in the next 1-10 years. Miscalibration was noted, with an under-prediction of risk seen particularly at higher predicted probabilities. Depending on how the model might be used, such under-prediction may be less of a concern; for example, if the model was being used to inform treatment changes only above a certain threshold of predicted risk. In this context, physicians could be confident that the true risk is at least at this threshold, if not higher. However, further studies are needed to explore the appropriate thresholds which maximise the model's clinical utility and cost-effectiveness, and to examine whether recalibration is possible in local settings.

The model may also be used to target falls prevention strategies at patients with the highest risk. These strategies might include multifactorial or exercise based interventions (49), or review of prescribed medications, with those likely to increase the risk of falls being considered for deprescribing (4, 18). Such medication reviews are increasingly being encouraged in routine clinical practice and the STRATIFY-Falls model may be useful for informing these (54). For example, in patients prescribed antihypertensive therapy, the model might be used alongside a cardiovascular risk prediction algorithm to compare the potential for benefit and harm from continued treatment prescription (26, 27, 55). For individuals with a high risk of falls but low risk of CVD, a physician might consider whether new or continued antihypertensive treatment is still appropriate. We examined the prevalence of this scenario in our model development population (Figure 5) and identified a significant number of individuals (11%) who would be classified in this way in terms of 10-year risk. More common were individuals with a low risk of falls but high risk of CVD (affecting 1 in 3 patients). For these patients, the model could be used by physicians to illustrate the minimal risk of harm for an individual, potentially improving uptake of, adherence to, and persistence with antihypertensive therapy, which is known to be poor currently (56).

## **Conclusions**

The STRATIFY-Falls prediction model helps to identify those at high risk of falls and could be used by physicians wishing to identify patients who might benefit from targeted fall prevention strategies, including multifactorial or exercise based interventions (49) and medication reviews. Used alongside other prediction tools such as those for cardiovascular risk, such a model could be valuable when used as part of a wider risk assessment for falls prevention.

## **Acknowledgements**

We thank Lucy Curtin for administrative support throughout the project. We thank Margaret Ogden, Prof Simon Griffin, and Prof John Gladman for their contributions as STRATIFY Investigators to the project. The Hospital Episode Statistics data used in this analysis are re-used with permission from NHS Digital who retain the copyright for those data. We thank the Office for National Statistics for providing data on mortality. The Office for National Statistics, and NHS Digital bear no responsibility for the analysis or interpretation of the data. Finally, we are very grateful to all those patients who permit their anonymised routine NHS data to be used for this approved research.

We thank the editors of The BMJ and the three reviewers for their helpful feedback that improved the article on revision.

## **Author contributions**

JPS conceived the project and wrote the protocol with FDRH, RJM, RS, RDR. CK and SLF extracted data for analysis. CK developed the model under supervision of JS and RS. LA validated the model under supervision of RDR and KIES. LA, KIES and RDR wrote the first draft of the manuscript. All authors revised the manuscript and approved the final version. JPS is the guarantor for this work and accepts full responsibility for the conduct of the study, had access to the data, and controlled the decision to publish. The corresponding author (JPS) attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

## **Funding source**

This research, JPS and CK were funded in whole, or in part, by the Wellcome Trust/Royal Society via a Sir Henry Dale Fellowship held by JPS (ref: 211182/Z/18/Z) and the National Institute for Health Research (NIHR) School for Primary Care (project 430) awarded to JPS.

JPS also receives funding via an NIHR Oxford Biomedical Research Centre (BRC) Senior Fellowship. RJMcM, is supported by an NIHR Senior Investigator award. FDRH acknowledges part support from the NIHR ARC Oxford Thames Valley, and the NIHR Oxford OUH BRC. KIES is funded by an NIHR SPCR Launching Fellowship. SLF was part funded by the NIHR BRC and NIHR ARC Oxford and Thames Valley. AB has received research funding from Astra-Zeneca, NIHR, BMA Medical Research Foundation and UKRI. RAP receives funding from the NIHR. AC is part funded by NIHR ARC Yorkshire & Humber and Health Data Research UK, an initiative funded by UK Research and Innovation Councils, National Institute for Health Research and the UK devolved administrations, and leading medical research charities. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care. For the purpose of Open Access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

### **Role of the funding source**

The sponsor and funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

### **Competing interests declaration**

All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) and declare: authors had financial support from the Wellcome Trust, Royal Society, and National Institute for Health Research for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

### **Ethics approval**

The study protocol was approved by CPRD's Independent Scientific Advisory Committee in February 2019 before obtaining the data relevant to the project (protocol given in the eAppendix in the Supplement). All data are fully anonymised so consent was not required. A project summary is published on the CPRD website (<https://www.cprd.com/isac>).

### **Transparency statement**

JPS affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as originally planned (and, if relevant, registered) have been explained.

### **Data sharing**

Data were obtained via a CPRD institutional licence. Requests for data sharing should be made directly to the CPRD. The algorithm is freely available for research use and can be downloaded from <https://process.innovation.ox.ac.uk/software/>. Code lists used to define variables included in the dataset are available at <https://github.com/jamessheppard48/STRATIFY-BP/tree/STRATIFY-Falls>

### **Dissemination to patient and public communities**

Findings from this study will be press released alongside publication of this manuscript. Social media (e.g., Twitter) will be used to draw attention to the work and stimulate debate about its findings. We will also publish a lay summary of our findings on our study website <https://www.phc.ox.ac.uk/research/stratified-treatments/studies/stratifying-treatments-in-the-multi-morbid-frail-elderly-stratify-antihypertensives> and make the underlying algorithms developed freely available for academic use here: [www.process.innovation.ox.ac.uk/software](https://process.innovation.ox.ac.uk/software).

Provenance and peer review: Not commissioned; externally peer reviewed.

### **Licence statement**

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which licence will apply to this Work are set out in our licence referred to above.

## References

1. Office for National Statistics. Living longer: how our population is changing and why it matters 2018 [Available from: [www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing](http://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/ageing).
2. Tinetti M, Speechley M, Ginter S. Risk factors for falls among elderly persons living in the community. *N Engl J Med*. 1988;319(26):1701-7.
3. Gale CR, Cyrus Cooper, Sayer AA. Prevalence and risk factors for falls in older men and women: The English Longitudinal Study of Ageing. *Age and Ageing*. 2016;45(6):789–94.
4. Gill, et al. Association of injurious falls with disability outcomes and nursing home admissions in community-living older persons. *Am J Epidemiol*. 2013;178:418-25.
5. Office for Health Improvement & Disparities. Public Health Outcomes Framework 2019-2020. 2021 [Available from: <https://fingertips.phe.org.uk/profile/public-health-outcomes-framework/data>
6. Whelton PK, Carey RM, Aronow WS, Casey DE, Jr., Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Hypertension*. 2018;71(6):e13-e115.
7. National Clinical Guideline Centre. Falls: Assessment and prevention of falls in older people [NICE guideline CG161]. London: Royal College of Physicians (UK). 2013.
8. White A, Tooth L, Peeters G. Fall risk factors in mid-age women: the Australian Longitudinal Study on Women's Health. *Am J Prev Med*. 2018;54(1):51-63.
9. Robbins A, Rubenstein L, Josephson K, Schulman B, Osterweil D, Fine G. Predictors of falls among elderly people: results of two population-based studies. *Arch Intern Med*. 1989;149:1628-33.
10. Gade G, Jørgensen M, Ryg J, et al. Predicting falls in community-dwelling older adults: a systematic review of prognostic models. *BMJ Open*. 2021;11(e044170).
11. de Vries M, Seppala L, Daams J, van de Glind E, Masud T, van der Velde N, et al. Fall-Risk-Increasing Drugs: A Systematic Review and Meta-Analysis: I. Cardiovascular Drugs. *J Am Med Dir Assoc*. 2018;19(4):371.e1-e9.
12. Seppala L, Wermelink A, de Vries M, Ploegmakers K, van de Glind E, Daams J, et al. Fall-Risk-Increasing Drugs: A Systematic Review and Meta-Analysis: II. Psychotropics. *J Am Med Dir Assoc*. 2018;19(4):371.e11-e17.
13. Seppala L, van de Glind E, Daams J, Ploegmakers K, de Vries M, Wermelink A, et al. Fall-Risk-Increasing Drugs: A Systematic Review and Meta-analysis: III. Others. *J Am Med Dir Assoc*. 2018;19(4):372.e1-e8.
14. Ettehad D, Emdin C, Kiran A, Anderson S, Callender T, Emberson J, et al. Blood pressure lowering for prevention of cardiovascular disease and death: a systematic review and meta-analysis. *Lancet*. 2016;387(10022):957-67.
15. Albasri A, Hattle M, Koshiares C, Dunnigan A, Paxton B, Fox S, et al. Association between antihypertensive treatment and adverse events: systematic review and meta-analysis. *BMJ (Clinical research ed)*. 2021;372:n189.
16. Leipzig R, Cumming R, Tinetti M. Drugs and falls in older people: A systematic review and meta-analysis: II Cardiac and analgesic drugs. *Journal of the American Geriatric Society*. 1999;47(1):40-50.
17. Tinetti M, Han L, Lee D, et al. Antihypertensive Medications and Serious Fall Injuries in a Nationally Representative Sample of Older Adults. *JAMA Intern Med*. 2014;174(4):588–95.
18. Reeve E, Gnjdic D, Long J, Hilmer S. A systematic review of the emerging definition of 'deprescribing' with network analysis: implications for future research and clinical practice. *British journal of clinical pharmacology*. 2015;80(6):1254-68.



19. Smith MI, de Lusignan S, Mullett D, Correa A, Tickner J, Jones S. Predicting Falls and When to Intervene in Older People: A Multilevel Logistical Regression Model and Cost Analysis. *PLOS ONE*. 2016;11(7):e0159365.
20. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data Resource Profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol*. 2015;44(3):827-36.
21. Wolf A, Dedman D, Campbell J, Booth H, Lunn D, Chapman J, et al. Data resource profile: Clinical Practice Research Datalink (CPRD) Aurum. *Int J Epidemiol*. 2019;48(6):1740-g.
22. 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 (Clinical research ed)*. 2008;336(7659):1475-82.
23. Wehner-Hewson N, Watts P, Buscombe R, Bourne N, Hewson D. Racial and Ethnic Differences in Falls Among Older Adults: a Systematic Review and Meta-analysis. *Journal of racial and ethnic health disparities*. 2021.
24. National Clinical Guideline Centre. Surveillance of the Falls in older people: assessing risk and prevention [NICE guideline CG161]. Appendix A: Summary of evidence review. London: Royal College of Physicians (UK) 2019.
25. Clegg A, Bates C, Young J, Ryan R, Nichols L, Ann Teale E, et al. Development and validation of an electronic frailty index using routine primary care electronic health record data. *Age Ageing*. 2016;45(3):353-60.
26. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ (Clinical research ed)*. 2017;357:j2099.
27. D'Agostino RB, Sr., Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. *Circulation*. 2008;117(6):743-53.
28. van Geloven N, Swanson SA, Ramspek CL, Luijken K, van Diepen M, Morris TP, et al. Prediction meets causal inference: the role of treatment in clinical prediction models. *European Journal of Epidemiology*. 2020;35(7):619-30.
29. Riley RD, Snell KI, Ensor J, Burke DL, Harrell FE, Jr., Moons KG, et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat Med*. 2019;38(7):1276-96.
30. Riley RD, Collins GS, Ensor J, Archer L, Booth S, Mozesder SI, et al. Minimum sample size calculations for external validation of a clinical prediction model with a time-to-event outcome. *Statistics in Medicine*. 2022;41(7):1280– 95.
31. White IR, Royston P. Imputing missing covariate values for the Cox model. *Statistics in Medicine*. 2009;28(15):1982-98.
32. Lau B, Lesko C. Missingness in the Setting of Competing Risks: from missing values to missing potential outcomes. *Curr Epidemiol Rep*. 2018;5(2):153–9.
33. DB R. Multiple Imputation for Nonresponse in Surveys. Wiley J, editor1987.
34. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*. 1999;94(446):496-509.
35. Feakins BG, McFadden EC, Farmer AJ, Stevens RJ. Standard and competing risk analysis of the effect of albuminuria on cardiovascular and cancer mortality in patients with type 2 diabetes mellitus. *Diagn Progn Res*. 2018;2:13.
36. Kawaguchi ES, Shen J, Li G, Suchard M. A Fast and Scalable Implementation Method for Competing Risks Data with the R Package fastcmprsk. *arXiv: Computation*. 2019.
37. Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol*. 1999;28(5):964-74.
38. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika*. 1982;69(1):239-41.

39. Graw F, Gerds T, Schumacher M. On pseudo-values for regression analysis in competing risks models. *Lifetime Data Anal.* 2009;15(2):241-55.
40. Andersen P, Perme M. Pseudo-observations in survival analysis. *Stat Methods Med Res.* 2010;19(1):71-99.
41. Royston P. Tools for Checking Calibration of a Cox Model in External Validation: Approach Based on Individual Event Probabilities. *The Stata Journal.* 2014;14(4):738-55.
42. Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Stat Med.* 2017;36:4391– 400.
43. Rahman MS, Ambler G, Choodari-Oskoei B, Omar RZ. Review and evaluation of performance measures for survival prediction models in external validation settings. *BMC medical research methodology.* 2017;17(1):60.
44. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *Bmc Medical Research Methodology.* 2009;9.
45. Riley RD, Ensor J, Snell KI, Debray TP, Altman DG, Moons KG, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ (Clinical research ed).* 2016;353:i3140.
46. Snell KI, Ensor J, Debray TP, Moons KG, Riley RD. Meta-analysis of prediction model performance across multiple studies: Which scale helps ensure between-study normality for the C-statistic and calibration measures? *Statistical methods in medical research.* 2017;962280217705678.
47. Debray TP, Damen JA, Riley RD, Snell K, Reitsma JB, Hooft L, et al. A framework for meta-analysis of prediction model studies with binary and time-to-event outcomes. 2018;28(9):2768-86.
48. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ (Clinical research ed).* 2016;352:i6.
49. Guirguis-Blake JM, Michael YL, Perdue LA, Coppola EL, Beil TL. Interventions to Prevent Falls in Older Adults: Updated Evidence Report and Systematic Review for the US Preventive Services Task Force. *Jama.* 2018;319(16):1705-16.
50. Lamb SE, Jørstad-Stein EC, Hauer K, Becker C. Development of a common outcome data set for fall injury prevention trials: the Prevention of Falls Network Europe consensus. *J Am Geriatr Soc.* 2005;53(9):1618-22.
51. Womack J, Murphy T, Bathulapalli H, Smith A, J B, Jarad S, et al. Serious Falls in Middle-Aged Veterans: Development and Validation of a Predictive Risk Model. *J Am Geriatr Soc.* 2020;68(12):2847-54.
52. Collins G, Reitsma J, Altman D, Moons K. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ (Clinical research ed).* 2015;350:g7594.
53. Banerjee A, Clementy N, Haguenoer K, Fauchier L, Lip GY. Prior history of falls and risk of outcomes in atrial fibrillation: the Loire Valley Atrial Fibrillation Project. *The American journal of medicine.* 2014;127(10):972-8.
54. Stewart D, et al. Structured medication reviews: origins, implementation, evidence, and prospects. *BJGP.* 2021;71(709):340-1.
55. Conroy RM, Pyorala K, Fitzgerald AP, Sans S, Menotti A, De Backer G, et al. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J.* 2003;24(11):987-1003.
56. Schulz M, Krueger K, Schuessel K, Friedland K, Laufs U, Mueller WE, et al. Medication adherence and persistence according to different antihypertensive drug classes: A retrospective cohort study of 255,500 patients. *International journal of cardiology.* 2016;220:668-76.



**Table 1: Descriptive statistics for model development and validation cohorts, in full cohorts and stratified by outcome type at 10 years. Values are n (%) unless otherwise stated.**

Variable	Development data			Validation data		
	Total (n=1 772 600)	Falls (n=62 691)	Competing event (n=181 731)	Total (n=3 805 366)	Falls (n=206 956)	Competing event (n=334 552)
Age, years – mean (SD)	59.4 (13.2)	73.6 (12.7)	74.3 (12.0)	58.6 (13.3)	72.8 (12.7)	73.1 (12.3)
Sex (Female)	921 853 (52%)	39 955 (64%)	91 676 (50%)	1 959 489 (52%)	134 945 (65.2%)	165 689 (49.5%)
Systolic blood pressure, mmHg	143.5 (11.9)	146.3 (12.7)	146.9 (12.8)	143.8 (12.3)	147.2 (13.2)	147.7 (13.3)
Diastolic blood pressure, mmHg	83.8 (9.6)	81.6 (10.0)	81.7 (10.0)	83.9 (9.8)	81.9 (10.2)	82.0 (10.3)
Cholesterol, mmol/L	5.3 (1.1)	5.2 (1.2)	5.2 (1.2)	5.5 (1.2)	5.4 (1.3)	5.4 (1.3)
Missing	868 461 (48.9%)	32 661 (52%)	104 094 (59%)	1 839 116 (48.3%)	109,708 (53.0%)	195,390 (58.4%)
Ethnicity						
White	734 149 (41%)	59 608 (95%)	105 077 (40.5%)	2 041 505 (54%)	194 311 (93.9%)	206 384 (61.7%)
Black	10 799 (0.6%)	339 (0.54%)	826 (0.45%)	115 279 (3%)	2 239 (1.1%)	4 019 (1.2%)
South Asian	14 799 (0.8%)	505 (0.81%)	991 (0.55%)	94 485 (3%)	2 449 (1.2%)	3 673 (1.1%)
Other	15 731 (0.9%)	587 (0.94%)	1 229 (0.68%)	832 614 (22%)	3 442 (1.7%)	21 458 (6.4%)
Missing	997 122 (56%)	1 652 (2.6%)	73 608 (57.8%)	721 483 (19%)	4 515 (2.2%)	99 018 (29.6%)
Deprivation Score						
IMD 1	420 765 (23.7%)	12 624 (20.2%)	35 529 (19.6%)	790 311 (20.8%)	41 786 (20.2%)	66 606 (19.9%)
IMD 2	406 775 (22.9%)	13 429 (21.4%)	39 652 (21.8%)	732 246 (19.2%)	41 820 (20.2%)	68 147 (20.4%)
IMD 3	376 765 (21.3%)	13 239 (21.1%)	39 279 (21.6%)	684 288 (18%)	40 665 (19.7%)	67 130 (20.1%)
IMD 4	313 595 (17.7%)	12 031 (19.2%)	35 183 (19.4%)	630 482 (16.6%)	40 383 (19.5%)	65 342 (19.5%)
IMD 5	254 700 (14.4%)	11 317 (18.1%)	31 909 (17.6%)	597 180 (15.7%)	42 141 (20.4%)	67 024 (20.0%)
Missing	0 (0%)	0 (0%)	0 (0%)	370 859 (9.7%)	161 (0.1%)	303 (0.1%)
Smoking status						
Non smoker	847 205 (48%)	29 500 (47.1%)	74 646 (41%)	1 475 708 (39%)	77 990 (37.7%)	109 249 (32.7%)
Ex-smoker	471 005 (27%)	17 440 (27.8%)	50 884 (28%)	1 236 061 (33%)	39 087 (18.9%)	75 081 (22.4%)

Smoker	363 440 (21%)	10 720 (17.1%)	38 478 (21.2%)	838 404 (22%)	66 836 (32.3%)	105 363 (31.5%)
Missing	90 950 (5%)	5 031 (8.0%)	17 905 (9.9%)	255 193 (7%)	23 043 (11.1%)	44 859 (13.4%)
Frailty index						
FI cont. (p50, IQR)	0.03 (0 to 0.08)	0.08 (0.06 to 0.14)	0.08 (0.06 to 0.14)	0.06 (0.03 to 0.08)	0.08 (0.06 to 0.17)	0.08 (0.06 to 0.17)
Alcohol						
Non drinker	289 472 (16%)	14 172 (22.6%)	37 568 (20.7%)	864 865 (23%)	59 364 (28.7%)	89 537 (26.8%)
Trivial drinker	488 289 (28%)	15 195 (24.2%)	42 645 (23.5%)	998 948 (26%)	47 088 (22.8%)	71 739 (21.4%)
Light drinker	239 732 (14%)	6 472 (10.3%)	18 863 (10.4%)	696 369 (18%)	26 635 (12.9%)	44 924 (13.4%)
Moderate drinker	179 102 (10%)	3 891 (6.2%)	12 926 (7.1%)	246 468 (7%)	9 378 (4.5%)	17 491 (5.2%)
Heavy drinker	22 760 (1.3%)	891 (1.4%)	2 336 (1.3%)	74 005 (2%)	5 124 (2.5%)	6 845 (2.1%)
Unknown amount	291 649 (16%)	9 962 (15.9%)	165 132 (14.4%)	237 464 (6%)	9 631 (4.7%)	12 117 (3.6%)
Missing	216 596 (15%)	12 108 (19.3%)	41 261 (22.7%)	687 247 (18%)	49 736 (24%)	91 899 (27.5%)
Risk Factors						
Previous falls	108 745 (6%)	10 514 (16.8%)	22 459 (12.4%)	140 886 (3.7%)	21 697 (10.5%)	25 124 (7.5%)
Memory problems	28 276 (1.6%)	3 860 (6.2%)	10 556 (5.8%)	99 264 (2.6%)	15 996 (7.7%)	28 636 (8.6%)
Mobility problems	20 425 (1.2%)	2 462 (3.9%)	7 347 (4.0%)	85 675 (2.3%)	13 999 (6.8%)	22 928 (6.9%)
Stroke	44 339 (2.5%)	4 320 (6.9%)	14 167 (7.8%)	111 462 (2.9%)	15 704 (7.6%)	26 703 (8%)
Multiple sclerosis	6 367 (0.4%)	300 (0.5%)	798 (0.4%)	11 328 (0.3%)	975 (0.5%)	1 373 (0.4%)
Antihypertensive drugs						
ACE inhibitors	219 506 (12%)	12 039 (19.2%)	38 096 (20.9%)	478 778 (13%)	38 867 (18.8%)	67 787 (20.3%)
Angiotensin II receptor blockers	59 075 (3%)	3 167 (5.1%)	7 628 (4.2%)	136 926 (4%)	11 018 (5.3%)	14 308 (4.3%)
Alpha blockers	34 338 (2%)	2 088 (3.3%)	6 794 (3.7%)	68 131 (2%)	6 335 (3.1%)	11 388 (3.4%)
Beta blockers	216 122 (12%)	10 885 (17.4%)	31 341 (17.3%)	461 329 (12%)	36 317 (17.6%)	59 019 (17.6%)
Calcium channel blockers	193 141 (11%)	11 570 (18.5%)	35 859 (19.7%)	426 151 (11%)	37 590 (18.2%)	63 764 (19.1%)
Diuretics	180 065 (10%)	10 706 (17.1%)	29 783 (16.4%)	397 980 (11%)	36 418 (17.6%)	55 934 (16.7%)
Other antihypertensives	10 784 (0.6%)	400 (0.8%)	1 594 (0.9%)	19 235 (1%)	1 437 (0.7%)	2 471 (0.7%)

Other drugs						
Opioids	553 344 (31%)	26 060 (41.6%)	69 496 (38.2%)	1 213 876 (32%)	84 108 (40.6%)	121 303 (36.3%)
Hypnotics, anxiolytics	376 885 (21%)	17 703 (28.2%)	48 636 (26.8%)	750 584 (20%)	52 854 (25.5%)	78 627 (23.5%)
Antidepressants	383 647 (21%)	17 159 (27.4%)	42 767 (23.5%)	793 690 (21%)	52 820 (25.5%)	71 452 (21.4%)
Anticholinergic medications	207 345 (11%)	11 085 (17.7%)	29 384 (16.2%)	388 513 (10%)	31 542 (15.2%)	46 255 (13.8%)
Follow up, years (p50, IQR)	6.2 (2.6 to 10)	4.3 (1.8 to 7.0)	3.7 (1.6 to 6.3)	6.7 (2.7 to 10)	4.3 (1.9 to 7.1)	3.8 (1.6 to 6.5)

**Table 2: Prediction model for falls. Values are sub-distribution hazard ratios and 95% confidence intervals.**

<b>Predictors</b>	<b>Full case analysis (n=358 207)</b>	<b>Multiple Imputation model (n=1 772 600)</b>
Age	30.1 (27.7 to 32.7)	60.46 (57.87 to 63.17)
Sex (Female)	1.32 (1.28 to 1.35)	1.25 (1.23 to 1.27)
Total cholesterol	1.55 (1.44 to 1.67)	1.48 (1.36 to 1.61)
<b>Ethnicity</b>		
White	Reference	Reference
Black	0.68 (0.59 to 0.79)	0.65 (0.58 to 0.74)
South Asian	0.67 (0.60 to 0.75)	0.68 (0.61 to 0.77)
Other	0.66 (0.59 to 0.74)	0.70 (0.63 to 0.78)
<b>Deprivation Score</b>		
IMD 1	Reference	Reference
IMD 2	1.05 (1.00 to 1.09)	1.04 (1.01 to 1.07)
IMD 3	1.06 (1.02 to 1.12)	1.07 (1.05 to 1.10)
IMD 4	1.14 (1.01 to 1.19)	1.18 (1.15 to 1.21)
IMD 5	1.23 (1.18 to 1.29)	1.35 (1.31 to 1.39)
<b>Smoking status</b>		
Non-smoker	Reference	Reference
Ex-smoker	1.06 (1.04 to 1.09)	1.12 (1.10 to 1.14)
Smoker	1.26 (1.22 to 1.31)	1.27 (1.24 to 1.30)
<b>Alcohol</b>		
Non-drinker	Reference	Reference
Trivial drinker	0.87 (0.84 to 0.90)	0.90 (0.85 to 0.95)
Light drinker	0.93 (0.89 to 0.98)	0.94 (0.88 to 1.00)
Moderate drinker	0.99 (0.94 to 1.05)	0.99 (0.93 to 1.06)
Heavy drinker	1.71 (1.55 to 1.87)	1.57 (1.28 to 1.93)
Unknown amount	0.97 (0.95 to 1.02)	0.93 (0.89 to 0.98)
Frailty index	1.11 (1.09 to 1.14)	1.22 (1.20 to 1.23)
<b>Risk Factors</b>		
Falls	1.40 (1.35 to 1.46)	1.32 (1.29 to 1.35)
Memory problems	1.25 (1.17 to 1.35)	1.17 (1.12 to 1.21)
Mobility problems	0.99 (0.93 to 1.07)	0.92 (0.87 to 0.98)
Stroke	1.28 (1.22 to 1.34)	1.14 (1.11 to 1.18)
Multiple sclerosis	1.48 (1.23 to 1.78)	1.71 (1.51 to 1.94)
<b>Antihypertensive drugs</b>		
ACE inhibitors	1.04 (1.01 to 1.07)	1.12 (1.10 to 1.14)
Angiotensin II receptor blockers	1.07 (1.02 to 1.12)	1.19 (1.15 to 1.23)
Alpha blockers	1.00 (0.95 to 1.06)	1.04 (1.02 to 1.06)
Beta blockers	0.97 (0.96 to 1.00)	1.07 (1.02 to 1.12)
Calcium channel blockers	0.99 (0.97 to 1.03)	1.08 (1.06 to 1.11)

Diuretics	0.98 (0.95 to 1.01)	1.07 (1.05 to 1.10)
Other antihypertensives	1.08 (0.97 to 1.21)	0.96 (0.88 to 1.04)
<b>Other drugs</b>		
Opioids	1.10 (1.07 to 1.13)	1.11 (1.08 to 1.13)
Hypnotics, anxiolytics	1.04 (1.00 to 1.07)	1.15 (1.13 to 1.18)
Antidepressants	1.14 (1.10 to 1.18)	1.16 (1.13 to 1.18)
Anticholinergic medications	1.11 (1.06 to 1.14)	1.03 (1.02 to 1.05)
<b>Variable transformations</b>		
Age = ((Age/100)^3) - 0.242		
Cholesterol = ((Cholesterol/10)^-0.5) - 1.381		
Frailty index = (Frailty index/0.1) - 0.576		



**Table 3: Predictive performance statistics of the falls prediction models on external validation in CPRD Aurum**

	1 year	5 years		10 years	
	Original model	Original model	Pseudo-value recalibration	Original model	Pseudo-value recalibration
<b>Observed/Expected</b>					
Pooled effect size (95% CI)	0.162 (0.158 to 0.166)	1.702 (1.674 to 1.730)	1.906 (1.874 to 1.939)	1.682 (1.657 to 1.707)	1.839 (1.811 to 1.865)
Prediction interval	0.090 to 0.289	1.116 to 2.586	1.246 to 2.915	1.139 to 2.484	1.284 to 2.638
Tau <sup>2</sup>	0.089 (0.080 to 0.099)	0.046 (0.042 to 0.052)	0.0479 (0.043 to 0.054)	0.038 (0.035 to 0.043)	0.0342 (0.031 to 0.038)
<b>C statistic</b>					
Pooled effect size (95% CI)	0.866 (0.862 to 0.869)	0.843 (0.841 to 0.844)	0.843 (0.841 to 0.844)	0.833 (0.832 to 0.835)	0.833 (0.831 to 0.835)
Prediction interval	0.794 to 0.915	0.789 to 0.881	0.789 to 0.881	0.789 to 0.870	0.789 to 0.870
Tau <sup>2</sup>	0.068 (0.056 to 0.083)	0.026 (0.023 to 0.030)	0.026 (0.023 to 0.030)	0.022 (0.019 to 0.025)	0.022 (0.019 to 0.025)
<b>D statistic</b>					
Pooled effect size (95% CI)	2.160 (1.987 to 2.333)	1.903 (1.754 to 2.051)	1.894 (1.746 to 2.042)	1.643 (1.515 to 1.771)	1.597 (1.472 to 1.721)
Prediction interval	1.99 to 2.33	1.75 to 2.05	1.75 to 2.04	1.51 to 1.77	1.47 to 1.72
Tau <sup>2</sup>	0.000 (0.000 to 0.039)	0.000 (0.000 to 0.023)	0.000 (0.000 to 0.022)	0.000 (0.0000 to 0.0168)	0.000 (0.000 to 0.016)
<b>Royston and Sauerbrei's <math>R_p^2</math></b>					
Range	0 to 86.0	28.0 to 91.4	25.9 to 91.4	21.3 to 91.4	21.6 to 91.4
Median (IQR)	58.1 (52.3 to 62.2)	47.4 (43.5 to 51.8)	47.3 (43.2 to 51.7)	39.9 (36.4 to 43.8)	38.6 (35.4 to 42.4)
Mean (SD)	56.5 (0.10)	47.9 (0.07)	47.7 (0.07)	40.8 (0.07)	39.4 (0.07)

## Box 1: Final model equations for predicting risk of falls at 1, 5 and 10 years

Final model equations, 5 and 10 years:

$$\ln\left(\frac{p_{5y}}{1-p_{5y}}\right) = \alpha_{5y} + (\beta_{5y} * LP) + (\gamma_{5y} * (LP) * \ln(LP))$$

$$\ln\left(\frac{p_{10y}}{1-p_{10y}}\right) = \alpha_{10y} + (\beta_{10y} * LP) + (\gamma_{10y} * (LP) * \ln(LP))$$

Final model equation for 1 year:

$$1 \text{ year risk} = 1 - (1 - CIF_{1y})^{\exp(LP)}$$

Where  $p_{5y}$  and  $p_{10y}$  are the predicted probabilities of a fall at 5 and 10 years respectively; and LP is the linear predictor from the original model, as shown in Table 2:

$$LP_i = \left(\left(\frac{Age}{100}\right)^3 - 0.242\right) * 4.102 + \begin{cases} 0.223 \text{ if female} \\ 0 \text{ if male} \end{cases} + \left(\left(\frac{TC}{10}\right)^{-0.5} - 1.381\right) * 0.393$$

$$+ \begin{cases} -0.425 \text{ if black ethnicity} \\ -0.381 \text{ if south asian ethnicity} \\ -0.352 \text{ if other ethnicity} \end{cases} + \begin{cases} 0.038 \text{ if IMD2} \\ 0.072 \text{ if IMD3} \\ 0.169 \text{ if IMD4} \\ 0.229 \text{ if IMD5} \end{cases} + \begin{cases} 0.114 \text{ if ex smoker} \\ 0.236 \text{ if current smoker} \end{cases}$$

$$+ \begin{cases} -0.105 \text{ if trivial drinker} \\ -0.065 \text{ if light drinker} \\ -0.009 \text{ if moderate drinker} \\ 0.451 \text{ if heavy drinker} \\ -0.068 \text{ if drinker (unkown quantity)} \end{cases} + \left(\left(\frac{FI}{0.1}\right) - 0.576\right) * 0.197$$

$$+ (0.275 \text{ if Previous falls}) + (0.155 \text{ if Memory problems})$$

$$+ (-0.08 \text{ if Mobility problems}) + (0.111 \text{ if Ace inhibitors})$$

$$+ (0.17 \text{ if Angiotensins}) + (0.082 \text{ if Calcium channel blockers})$$

$$+ (0.071 \text{ if Diuretics}) + (0.068 \text{ if Beta blockers}) + (0.041 \text{ if Alpha blockers})$$

$$+ (-0.045 \text{ if Other hypertensives}) + (0.101 \text{ if Opioids})$$

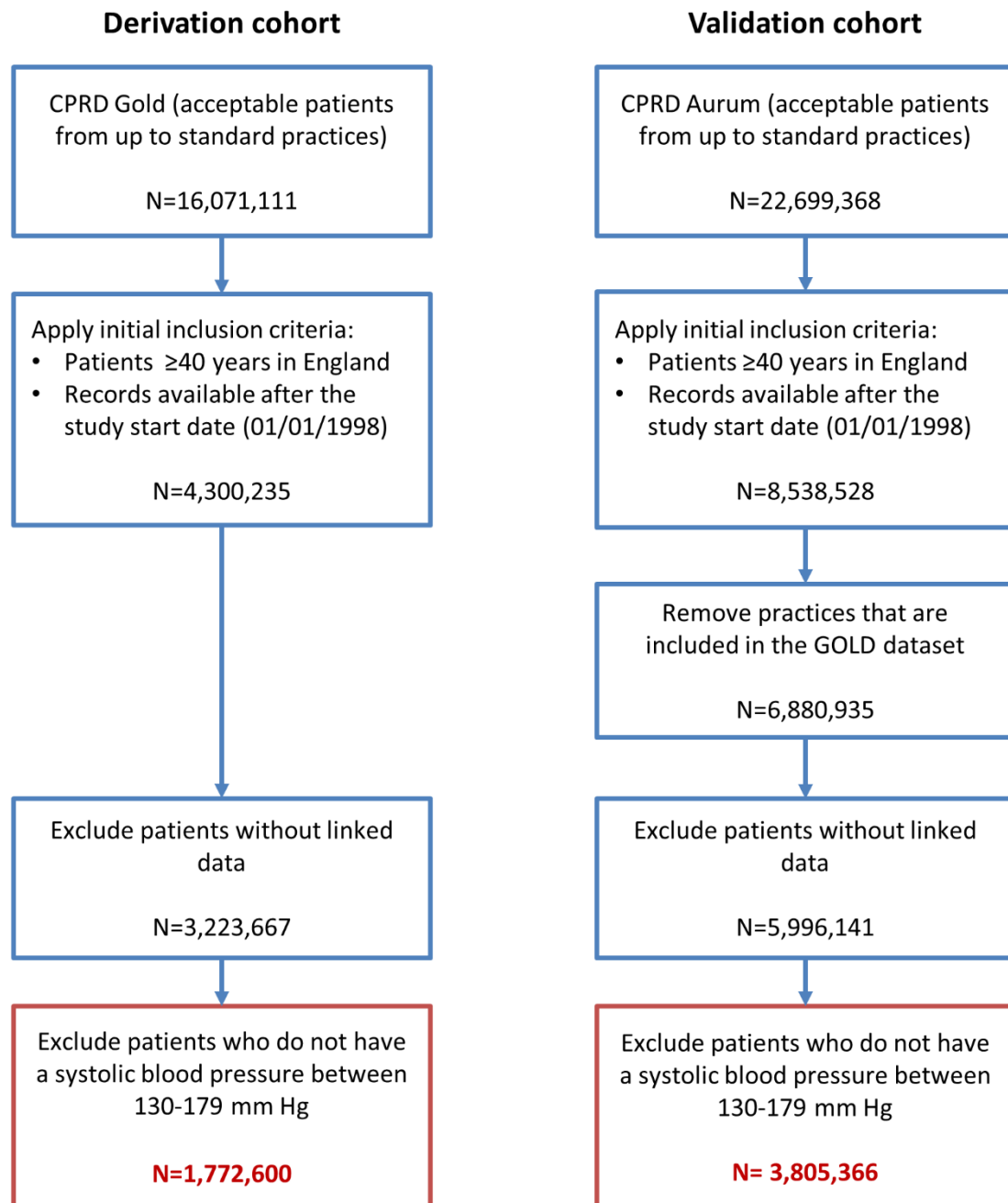
$$+ (0.142 \text{ if Hypnotics/Benzo}) + (0.146 \text{ if Antidepressants})$$

$$+ (0.034 \text{ if Anticholinergic}) + (0.133 \text{ if History of stroke})$$

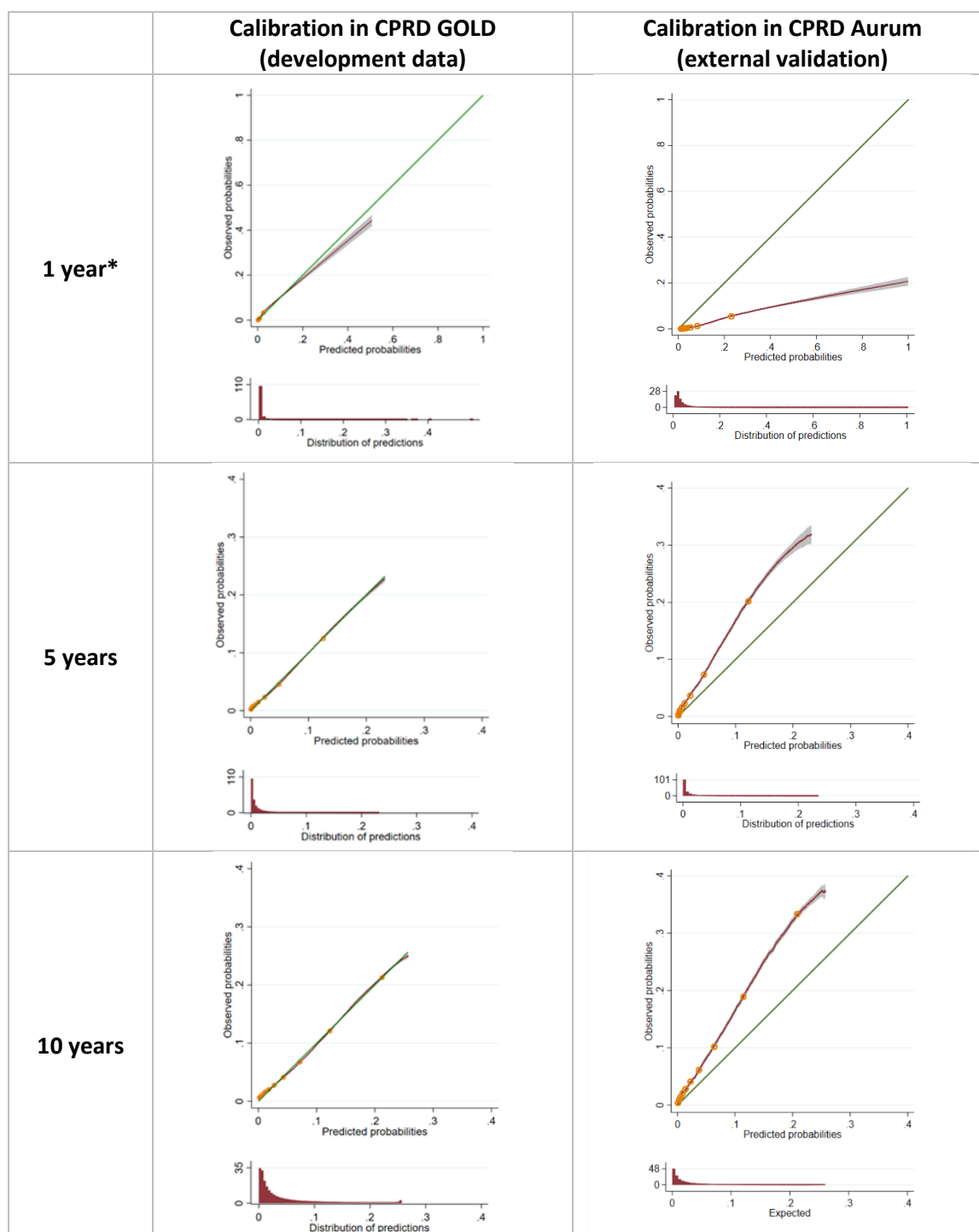
$$+ (0.537 \text{ if History of MS})$$

- ln is the natural logarithm
- IMD2-IMD5 refer to indices of multiple deprivation
- Age is measured in years
- TC refers to Total Cholesterol
- FI refers to Frailty Index
- The full algorithm code (including the  $\alpha, \beta, \gamma$  and CIF values) is freely available for research use and can be downloaded at [www.process.innovation.ox.ac.uk/software](http://www.process.innovation.ox.ac.uk/software) (link will be publicly available upon publication).

**Figure 1: Flow of study participants**



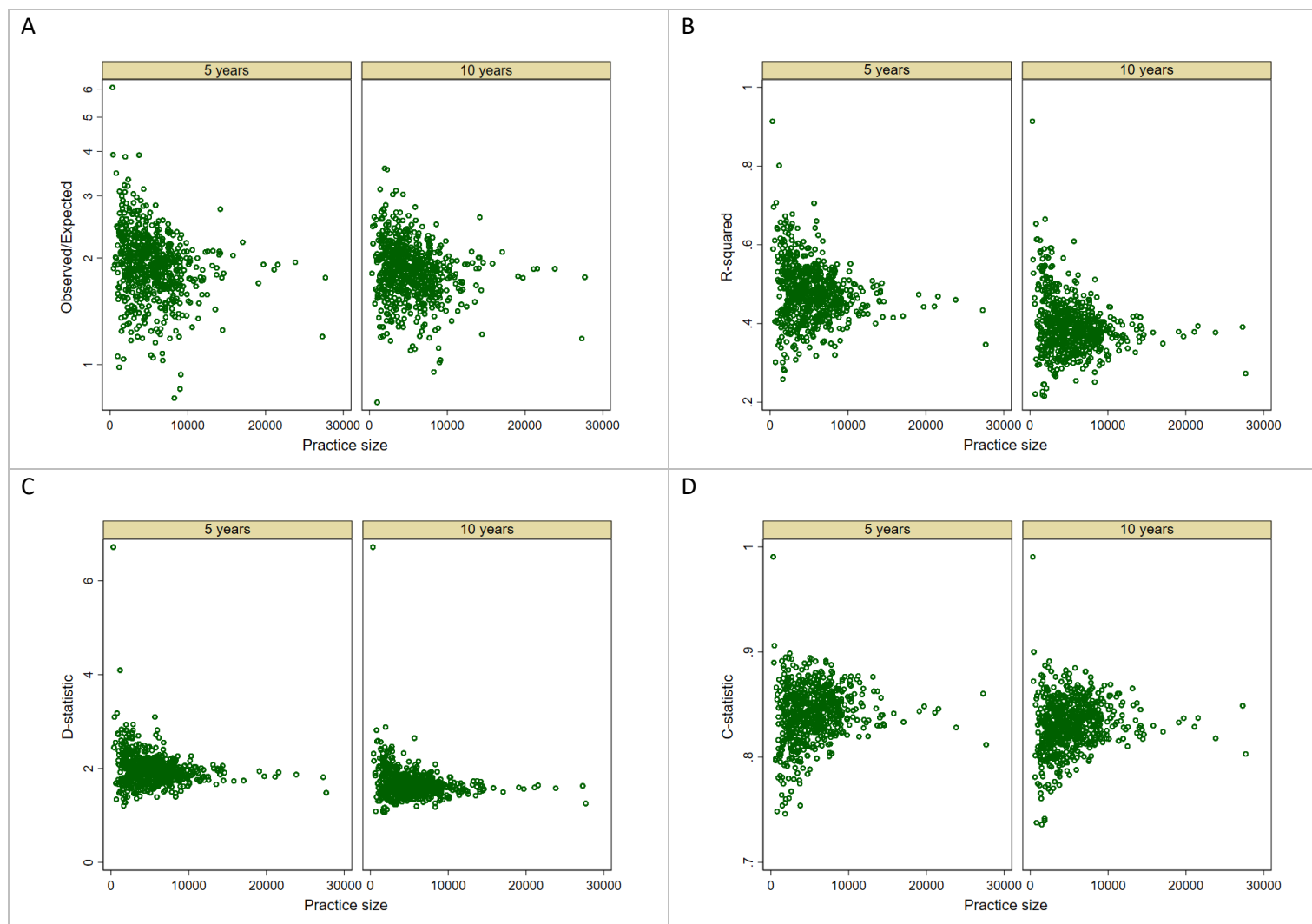
**Figure 2: Calibration curves for the apparent performance of the final STRATIFY-Falls model in CPRD GOLD, and the calibration on external validation in CPRD Aurum. Groups represent tenths of the linear predictor, as created between deciles. Histograms show the distribution of predicted probabilities.**



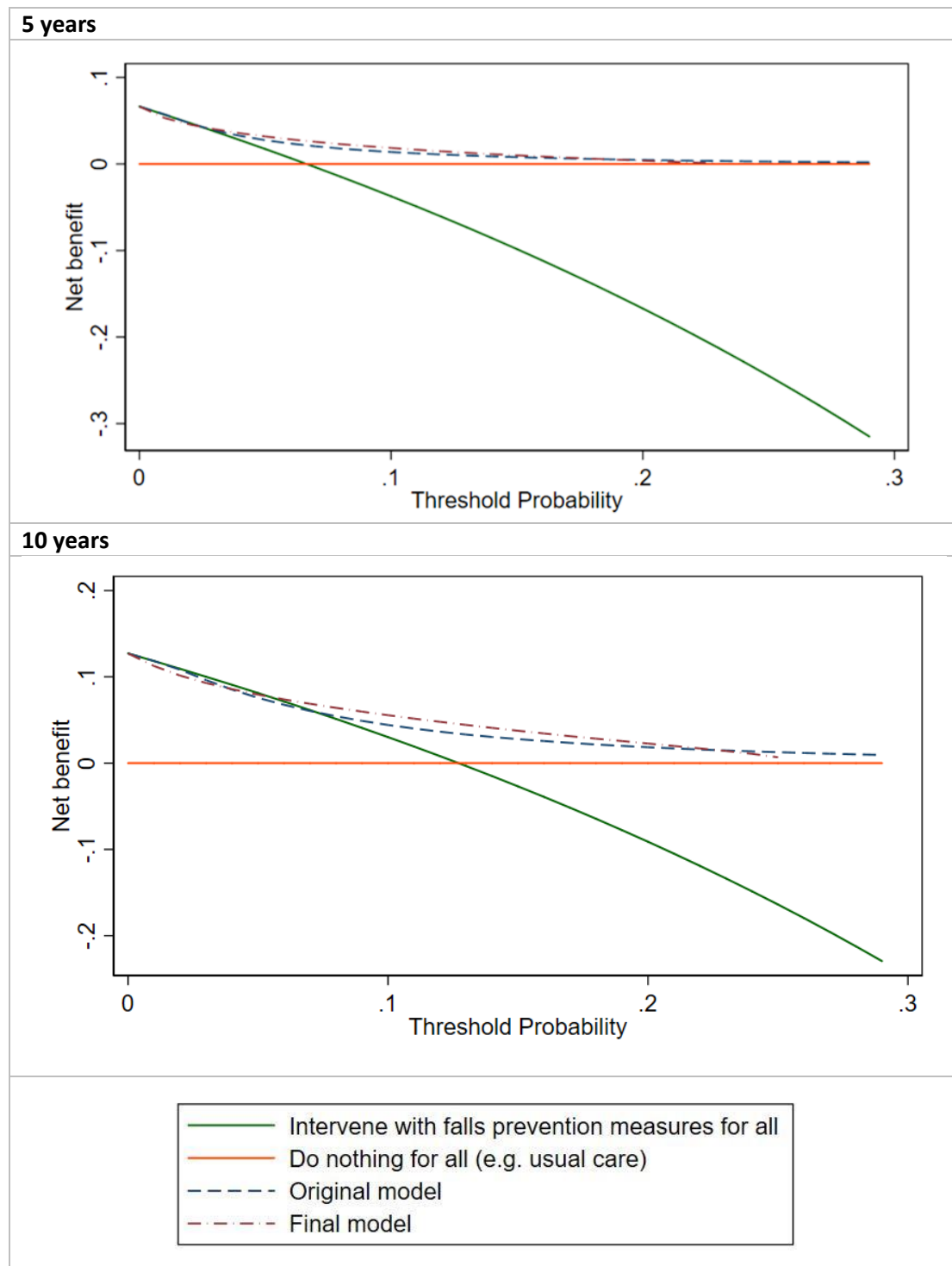
\*Note that the 1-year model is not recalibrated to pseudo-values in the development data

Figure 3: Performance variability of the STRATIFY-FALLS model on external validation across GP practices, with panels (A)

Observed/Expected, (B)  $R_D^2$ , (C) D-statistic, and (D) C-statistic



**Figure 4: Decision curve analysis, showing the Net Benefit of using prediction models across different threshold probabilities for assigning treatment**



**Figure 5: Comparison of 10-year CVD risk (Qrisk2) and fall risk in the GOLD dataset.**

**High risk for both conditions was defined as a risk above 10%**

