

Title: The utility of long-term blood pressure variability for cardiovascular risk prediction in primary care

Short title: Blood pressure variability and risk

Authors:

Sarah L STEVENS, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK. Richard J MCMANUS, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK. Richard J STEVENS, Nuffield Department of Primary Care Health Sciences, University of Oxford, UK.

Funding: This paper presents independent research funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR). Richard McManus was funded by an NIHR Research Professorship and by NIHR Oxford CLAHRC during the study. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Conflicts of interest: RM has received BP monitoring equipment for research purposes from Omron.

Corresponding author: Ms Sarah Stevens, Nuffield Department of Primary Care Health Sciences, Radcliffe Observatory Quarter, Woodstock Road, Oxford, OX2 6GG.
sarah.stevens@phc.ox.ac.uk.

Word count: 3636 **Tables:** 3. **Figures:** 3. **Supplementary files:** 2

ABSTRACT

Objectives: Blood pressure is a long-established risk factor for cardiovascular disease.

Systolic blood pressure is used in all widely-used cardiovascular risk scores for clinical decision-making. Recently, within-person blood pressure variability has been shown to be a major predictor of cardiovascular disease. We investigated whether cardiovascular risk scores could be improved by incorporating blood pressure variability with standard risk factors.

Methods: We used cohort data on patients aged 40 to 74 on 1/1/2005, from English general practices contributing to the Clinical Practice Research Datalink, a research database derived from electronic health records. Data were linked to hospital episodes and mortality data. Systolic blood pressure variability independent of the mean was calculated across up to 6 clinic visits. We divided data geographically into derivation and validation datasets. In the derivation dataset we developed a reference model, incorporating risk factors used in previous scores and an index model, incorporating the same factors plus blood pressure variability. We calculated model validation statistics in the validation dataset including calibration ratio and c-statistic.

Results: In the derivation dataset, blood pressure variability was associated with cardiovascular disease, independently of other risk factors ($p=0.005$). However, in the validation dataset, both models had similar c-statistic (0.7415 and 0.7419 respectively), R^2 (31.8 and 32.0 respectively) and calibration ratio (0.938 and 0.940 respectively).

Conclusions: The association of blood pressure variability with cardiovascular disease is statistically significant in a large dataset but does not substantially improve the performance of a cardiovascular risk score.

Key words: blood pressure; risk; risk assessment; coronary disease; stroke

INTRODUCTION

Over recent years, variability in biological factors over time, and possible associations with outcomes, have been studied in many clinical fields, particularly where only mean or average levels of biological factors have been considered historically, for example in the case of cholesterol and cardiovascular disease (CVD),(1,2) or blood glucose and mortality in critical care settings.(3) One of the most well studied is that of blood pressure (BP) variability, which has been shown to be associated with cardiovascular and mortality outcomes in the long, mid and short-term, independently of mean BP.(4)

As an established risk factor for CVD, a measure of mean BP is included in many CVD risk scores, such as the Framingham,(5) QRISK2(6) and SCORE(7) equations. Such scores are routinely used as part of guideline recommended primary prevention of CVD,(8) to identify those at highest risk in order to target treatment. However, they have not historically incorporated measures of BP variability and the utility of BP variability in CVD risk prediction has been little studied. Previous studies may have been optimistic in their assessments because risk score performance was evaluated in the same dataset as that used to develop the score and may not be applicable to a UK primary care population.(9,10) The recently published QRISK3 score only assessed BP variability over two measurements and included patients taking antihypertensive medication, which may have confounded the relationship between BP variability and CVD outcomes.(11)

We therefore aimed to establish whether including a measure of long-term BP variability in a CVD risk score would improve its accuracy in terms of discrimination and calibration, avoiding the short-comings of previous work.

METHODS

Study design

We studied an open cohort of men and women aged 40 to 74 on 1st January 2005, registered for at least two years at English general practices contributing to the Clinical Practice Research Datalink (CPRD), with linked Index of Multiple Deprivation (IMD), ONS mortality and HES Inpatient data. Based on our pre-specified study population, eligible patients entered the study on the latest of the following dates: 1st January 2005, date of current registration with the practice plus two years, the practice up-to-standard date (the date after which no significant gaps in reporting occurred, as defined by CPRD) plus two years and coverage start date for linked data plus two years. However, owing to fewer than expected participants having the required number of repeat BP measurements (Table S1, Online Supplement, please see <http://hyper.ahajournals.org>), the minimum prior registration period was extended from two years to five years. This increased the number of available BP readings (Table S2), without the drawback of other alternatives considered (for example including those on antihypertensive treatment, which may have confounded the relationship between BP variability and CVD outcomes due to, for example, changes in medication adherence(12,13)).

Patients exited the study on the earliest of the following dates: date of death, transfer out date, last practice data upload date, coverage end date for linked data or index date plus ten years. Patients were excluded if they had any history of antihypertensive medication use or CVD prior to study entry or were taking statins in the year prior to study entry. We restricted analysis to a random sample of 200,000 eligible individuals in accordance with our

sample size calculations (see the Extended Methods section of the Online Supplement) and ethical approval given for this study

The primary outcome was first CVD event defined as a composite of myocardial infarction, coronary and ischaemic heart disease, angina, cerebrovascular and haemorrhagic stroke events and cause-specific mortality. Blood pressure mean and variability were calculated from a maximum of six BP measurements occurring at least one month apart prior to study entry, up to a maximum of 10 years.

The full study protocol was approved by the Independent Scientific Advisory Committee (ISAC) to the MHRA (protocol number 16_034R) and was made available during the peer review process. Ethical approval for observational research using the CPRD with approval from ISAC has been granted by a National Research Ethics Service committee (Trent MultiResearch Ethics Committee, REC reference number 05/MRE04/87). Extended methods, including code lists and details of data cleaning processes are provided in the Online Supplement.

Statistical analysis

Model derivation

The data was split into derivation and validation subsample by practice region; broadly South (South West, South East and London) for derivation and North (North, Midlands and East Anglia) for validation. In the derivation subset, a risk prediction model for CVD events was derived using Weibull parametric survival models. The following variables (included in previously developed scores such as Framingham(5) and QRISK2(6)) were included in the reference model without selection: age, gender, mean systolic blood pressure (mm Hg),

total cholesterol to high density lipoprotein cholesterol ratio (mmol/L), ethnicity, index of multiple deprivation (IMD) score, family history of angina or myocardial infarction or stroke, smoking status (never, ex, current smoker), body mass index (kg/ m²), history of diabetes (Type 1 or 2), history of chronic kidney disease (Stage 3-5), history of rheumatoid arthritis, history of atrial fibrillation and history of left ventricular hypertrophy. Interaction terms with age and fractional polynomial terms were included in the model if they were significant at the 10% level. Terms for BP variability (standard deviation, coefficient of variation, variation independent of the mean and average real variability) were added individually to the reference model. The term with the largest standardised hazard ratio was retained in the model and taken forward to validation (variability model).

Model validation

Validation of the reference and variability models was carried out in the derivation and validation datasets. Predicted risk was calculated from the models and observed risk was calculated from Kaplan-Meier estimates. The following model validation statistics were calculated for both models: E/O calibration statistic (ratio of expected (E) to observed (O) number of events),(14) c statistic for discrimination (Harrell's version, to allow for censored survival data),(15) Royston's R-squared statistic (transformation of D statistic)(16) and net reclassification improvement/ integrated discrimination improvement statistics.(10,17)Plots of the observed risk of events against predicted risk in each decile of predicted risk for men and women were also created. Due to the volume of data, it was computationally infeasible to adjust models for optimism using bootstrap methods.(18)

Missing data

Missing data were handled differently depending on the variable in question. For diagnoses, absence of a relevant READ/ ICD-10 code was assumed to reflect absence of disease.

Patients with missing ethnicity and smoking status records were assumed to be White and non-smokers respectively. Continuous variables were imputed using multiple imputation by chained equations,(19,20) including the event indicator and Nelson-Aalen estimate of the cumulative hazard in the imputation model.(21,22) Eighty imputed datasets were created for the derivation and validation subsets separately.

RESULTS

In total 863,880 eligible individuals were identified from whom 200,000 were randomly selected. The final study dataset included 199819 patients after removal of 181 people with missing IMD data (Figure 1). The derivation dataset included 102996 patients and 663492 person-years of follow-up. The validation dataset included 96823 patients with a total of 646684 person-years of follow-up. The characteristics of included individuals are given in Table 1. In total 79.6% of patients had at least one BP measurement recorded and 52.8% of patients had at least two BP measurements recorded, allowing calculation of a measure of BP variability (Table 1).

The final reference model is given in Table S3. Due to the low prevalence of left-ventricular hypertrophy, this was removed from the model. Interaction terms were included for age with total cholesterol to high density lipoprotein cholesterol ratio, smoking status and mean BP but no fractional polynomial terms were included. All four measures of long-term BP variability were associated with CVD when adjusting for variables in the reference model. BP variability independent of mean was included in the final variability model (Table 2) as this

had the largest standardized hazard ratio of the variability measures studied, although confidence intervals overlapped (Table S4).

Model validation

In internal validation in the derivation dataset, the performance of the reference and variability models was similar (Table S5, Figures S1 and S2). This was also the case in external validation in the validation dataset (Table 3). Both models had good discrimination with c-statistics >0.74 . Both models explained approximately 32% of the variation in risk and underestimated risk on average (expected divided by observed risk ratios were 0.94).

Underestimation was greater in those at higher risk and in men (Figure 2 and Figure 3).

The net reclassification improvement indicated that classification across the 10% risk threshold was non-significantly worse in the variability model, with a net change of -0.02% (equivalent to 20 people out of 96,823 in the validation cohort being incorrectly reclassified). The integrated discrimination improvement index indicated a significant but very small improvement.

Sensitivity and subgroup analyses

To assess to what extent missing data may have influenced the results, we repeated the model development and internal validation using only two repeat BP measurements and found that estimates of the association of BP variability with CVD were diluted across all measures of variability. In this analysis, coefficient of variation was included in the variability model but as in the main analysis, the reference and variability models performed similarly in internal validation (Table S6). In post-hoc subgroup analysis we found that the reference

and variability models performed similarly in external validation in both patients with missing BP data and those with fully observed BP readings (Table S7).

DISCUSSION

This study has shown that in routinely collected primary care data from England, blood pressure variability is significantly associated with 10-year cardiovascular risk, over and above traditional risk factors. However, including a term for BP variability in a risk score for CVD did not improve the accuracy of the risk score in terms of discrimination, calibration or proportion of people correctly classified into risk groups.

Strengths

This analysis included data from nearly 200,000 individuals with over 9500 events and would therefore be large enough to detect any meaningful differences in risk score performance. Model performance was similar to published performance estimates for the recommended risk score in the UK (QRISK2) and the previously recommended NICE modified Framingham equation.(23)

By using routinely collected data we have examined the real-world utility of BP variability as a novel risk factor for CVD risk prediction, whereas previous studies using trial(24) or prospective study data(9) may not be generalizable to primary care due to differences caused by trial or study selection.(25) Furthermore, the CPRD is broadly representative of the UK population,(26) and the study population was carefully selected to limit the potential for confounding. Previous studies have failed to exclude patients on treatment or to separate the measurement and follow-up periods.(24) The recently published QRISK3 algorithm also failed to account for treatment during the measurement period for BP

variability,(11) and hence these studies may be subject to confounding between BP variability, treatment and outcomes.(12,27)

Limitations

We encountered large amounts of missing data, particularly for blood pressure, and it is unclear to what extent this may have affected our results. However, we found that results were similar regardless of whether BP variability was measured over two or six readings. The lack of repeat BP measurements also represents a finding in itself, showing that regardless of model performance, the utility of BP variability is limited in practice.

Due to the volume of missing data, a large number of imputed datasets were required. We originally intended to carry out bootstrap validation analysis to estimate model optimism but this was computationally and practically infeasible when coupled with imputation. However, optimism is a particular problem in small samples where model over-fitting is more likely.(15) It may be reasonable to assume that optimism in the reference and variability models would be similar given they differ by only a single covariate, and relatively low given the large sample size relative to the number of included variables.

We chose not to test four separate models (one for each measure of BP variability) in our validation dataset, partly because of computational and time limitations. However, since the standardized hazard ratios for each measure of variability were similar and confidence intervals overlapped considerably, we are confident that results would have been similar had other measures of variability been considered. Our sensitivity analysis considering BP variability measured by coefficient of variation over two readings supports this.

The risk scores developed may not be generalizable to other populations, such as people under the age of 40, or those taking antihypertensive medications. Other risk scores, for example QRISK2 which was developed in patients aged 25 and over and explicitly includes a term for treated hypertension,(6) may be more applicable to such patients. Further validations in distinct cohorts would be required to confirm the performance of the derived risk scores.

Comparisons with existing literature

The estimated association of BP variability on CVD outcomes in this study is at the lower end of the 95% confidence interval for the estimated association calculated from our previous meta-analysis.(4) This is to be expected due to the potential for measurement error in BP readings available in CPRD (compared to bespoke studies) and this may have attenuated the blood pressure model parameters.

Studies assessing the added predictive value of risk factors are common, yet traditional methods of model assessment can be conservative.(10) It is therefore unsurprising that the variability model did not predict CVD more accurately according to traditional measures of discrimination and calibration. Although no improvement was observed even when using newer more sensitive measures of reclassification (net reclassification improvement or integrated discrimination improvement),(10) such a result is not uncommon and has been observed when assessing the incremental value of many risk factors including: subclinical hyperthyroidism,(28) carotid intima-media thickness,(29) and glycated haemoglobin.(30)

The recently published QRISK3 algorithm included a measure of BP standard deviation.(11) There were several methodological differences between the QRISK3 study and this study which could have driven contrasting conclusions. For example, the QRISK3 study included

patients on antihypertensive treatment and calculated BP variability from a minimum of only two repeat BP readings. However, despite these differences, the authors similarly found little additional benefit of including a measure of BP variability on overall risk score performance.

Implications for research and practice

This research shows that there is no benefit to updating existing cardiovascular risk scores used in primary care to include measures of long-term BP variability. Indeed, the large amount of missing data in this study indicates that, even if BP variability measures do have added predictive value in certain sub-populations, they would be of little practical use as the data required is simply not collected.

Future research could focus on the added predictive ability of BP variability over the medium and short-term, as measured by home and ambulatory BP, but would be similarly limited in generalisability unless patients undergoing out-of-office measurement included those other than with suspected hypertension only. Blood pressure variability may be more useful as a predictor in other settings (e.g. hospitals) where BP monitoring is more systematic and routine, but this has yet to be explored. Research in these settings must give careful consideration to the effect of treatments on BP variability and on outcomes (for example in dialysis patients who may experience intra-dialysis hypertension(31)) to ensure appropriate controls for confounding effects.

Systematic review evidence suggests that BP variability may well be more predictive of stroke events than coronary events ,(4) and future research could assess the added predictive ability of BP variability for stroke events specifically. However this may be of less

use clinically since current guidelines for primary prevention of CVD only advocate risk assessment for CVD as a composite end-point.(8)

The challenge of missing data highlights the limitation of using routinely collected data for some studies. Researchers using EHRs to study variability in BP or other biological factors should carefully consider the population to be studied to ensure adequate numbers of repeat measurements.

In conclusion, although long-term BP variability is an independent predictor of CVD, adding BP variability to a risk score including several traditional risk factors did not improve either the calibration or discrimination of the score. The limited number of repeat BP measurements taken in primary care suggests that measurement of BP variability may only be possible in select groups of patients or in other settings in practice.

ACKNOWLEDGEMENTS

None

REFERENCES

1. Perera R, McFadden E, McLellan J, Lung T, Clarke P, Pérez T, et al. Optimal strategies for monitoring lipid levels in patients at risk or with cardiovascular disease: a systematic review with statistical and cost-effectiveness modelling. *Health Technol Assess (Rockv)*. 2015 Dec;19(100):1–402.
2. Bangalore S, Breazna A, DeMicco DA, Wun C-C, Messerli FH. Visit-to-Visit Low-Density Lipoprotein Cholesterol Variability and Risk of Cardiovascular Outcomes: Insights From the TNT Trial. *J Am Coll Cardiol*. 2015;65(15):1539–48.
3. Egi M, Bellomo R, Reade MC. Is reducing variability of blood glucose the real but hidden target of intensive insulin therapy? *Crit Care*. 2009;13(2):302.
4. Stevens SL, Wood S, Koshiaris C, Law K, Glasziou P, Stevens RJ, et al. Blood pressure variability and cardiovascular disease: systematic review and meta-analysis. *BMJ*. 2016;(354):i4098.
5. Anderson KM, Odell PM, Wilson PWF, Kannel WB. Cardiovascular disease risk profiles. *Am Heart J*. 1991;121:293–8.
6. 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(7659):1475–82.
7. Conroy RM, Pyörälä 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.

8. National Institute for Health and Care Excellence. Lipid modification: cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease (NICE CG181). National Institute for Health and Care Excellence; 2014.
9. Palatini P, Reboldi G, Beilin LJ, Casiglia E, Eguchi K, Imai Y, et al. Added predictive value of night-time blood pressure variability for cardiovascular events and mortality: the Ambulatory Blood Pressure-International Study. *Hypertension*. 2014;64(3):487–93.
10. Pencina MJ, D’Agostino RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008 Jan 30;27(2):157-72-12.
11. 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*. 2017;357.
12. Muntner P, Levitan EB, Joyce C, Holt E, Mann D, Oparil S, et al. Association Between Antihypertensive Medication Adherence and Visit-to-Visit Variability of Blood Pressure. *J Clin Hypertens*. 2013;15(2):112–7.
13. Hong K, Muntner P, Kronish I, Shilane D, Chang TI. Medication adherence and visit-to-visit variability of systolic blood pressure in African Americans with chronic kidney disease in the AASK trial. *J Hum Hypertens*. 2015;(Apr 2):1–6.
14. Viallon V, Ragusa S, Clavel-Chapelon F, Bénichou J. How to evaluate the calibration of a disease risk prediction tool. *Stat Med*. 2009 Mar 15;28(6):901–16.

15. Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15(4):361–87.
16. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. *Stat Med*. 2004 Mar 15;23(5):723–48.
17. Pencina MJ, D’Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011 Jan 15;30(1):11–21.
18. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol*. 2001 Aug;54(8):774–81.
19. White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med*. 2011;30(4):377–99.
20. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009 Jan;338:b2393.
21. Moons KGM, Donders RART, Stijnen T, Harrell Jr. FE. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol*. 2006;59(10):1092–101.
22. White IR, Royston P. Imputing missing covariate values for the Cox model. *Stat Med*. 2009;28(15):1982–98.

23. Collins GS, Altman DG. Predicting the 10 year risk of cardiovascular disease in the United Kingdom: independent and external validation of an updated version of QRISK2. *BMJ*. 2012;344(7867):e4181.
24. Blacher J, Safar ME, Ly C, Szabo de Edelenyi F, Hercberg S, Galan P. Blood pressure variability: cardiovascular risk integrator or independent risk factor? *J Hum Hypertens*. 2015 Feb;29(2):122–6.
25. Kho ME, Duffett M, Willison DJ, Cook DJ, Brouwers MC. Written informed consent and selection bias in observational studies using medical records: systematic review. *BMJ*. 2009 Mar 12;338:b866.
26. 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 Jun 6;44(3):827–36.
27. Webb AJ, Fischer U, Mehta Z, Rothwell PM. Effects of antihypertensive-drug class on interindividual variation in blood pressure and risk of stroke: a systematic review and meta-analysis. *Lancet*. 2010;375(9718):906–15.
28. Kim TH, Choi HS, Bae JC, Moon JH, Kim H-K, Choi SH, et al. Subclinical hypothyroidism in addition to common risk scores for prediction of cardiovascular disease: a 10-year community-based cohort study. *Eur J Endocrinol*. 2014 Nov;171(5):649–57.
29. Bots ML, Groenewegen KA, Anderson TJ, Britton AR, Dekker JM, Engström G, et al. Common carotid intima-media thickness measurements do not improve cardiovascular risk prediction in individuals with elevated blood pressure: the USE-IMT collaboration. *Hypertension*. 2014 Jun;63(6):1173–81.

30. Di Angelantonio E, Gao P, Khan H, Butterworth AS, Wormser D, Kaptoge S, et al.
Glycated Hemoglobin Measurement and Prediction of Cardiovascular Disease. JAMA.
2014 Mar 26;311(12):1225.
31. Inrig JK, Patel UD, Toto RD, Szczech LA. Association of blood pressure increases during
hemodialysis with 2-year mortality in incident hemodialysis patients: a secondary
analysis of the Dialysis Morbidity and Mortality Wave 2 Study. Am J Kidney Dis. 2009
Nov;54(5):881–90.

Table 1: Baseline characteristics of 199,819 patients in the derivation and validation cohorts

	Derivation cohort	Validation cohort
	(N=102,996)	(N=96,823)
Variable	Mean (SD)/ N (%)	Mean (SD)/ N (%)
Female gender	51,150 (49.7)	47,332 (47.9)
Age (years)	53.9 (9.08)	53.3 (9.00)
CVD event outcome	4,562 (4.4)	5,105 (5.3)
Hypertension	812 (0.8)	938 (1.0)
Atrial fibrillation	189 (0.2)	169 (0.2)
Family history of CHD	14,836 (14.4)	18,207 (18.8)
Family history of stroke	3,665 (3.56)	3,797 (3.8)
Chronic kidney disease (Stage 3 to 5)	279 (0.3)	165 (0.2)
Diabetes (Type 1 or 2)	849 (0.8)	802 (0.8)
Rheumatoid arthritis	553 (0.5)	590 (0.6)
Left ventricular hypertrophy	15 (0.0)	24 (0.0)
Body mass index (kg/ m²)	26.7 (5.03)	26.9 (5.06)
	(N=51,649)	(N=44,774)
Total cholesterol (mmol/L)	5.53 (0.99)	5.59 (1.00)
	(N=29,896)	(N=25,467)
High density lipoprotein (mmol/L)	1.51 (0.46)	1.50 (0.43)
	(N=23,897)	(N=18,539)

Mean blood pressure (across all readings, mm Hg)	129.3 (14.5) (N=82,007)	130.4 (14.4) (N=76,980)
Blood pressure variability (across two or more readings)	(N=53,730 for all)	(N= 51,799 for all)
Standard deviation	9.42 (5.96)	9.64 (5.98)
Coefficient of variation	7.28 (4.52)	7.38 (4.49)
Average real variability	11.79 (8.19)	11.98 (8.21)
Variation independent of mean	9.44 (5.87)	9.65 (5.87)
Smoking status		
Non-smoker	40,311 (39.1)	37,923 (39.2)
Ex-smoker	16,582 (16.1)	14,323 (14.8)
Current smoker	20,506 (19.9)	21,138 (21.8)
Missing (assumed non-smoker)	25,597 (24.9)	23,439 (24.2)
Ethnicity		
White	42,992 (41.7)	43,006 (44.4)
Missing (assumed white)	43,989 (42.7)	38,000 (39.3)
Index of multiple deprivation		
Bottom quintile (least deprived)	33,992 (33.0)	22,569 (23.3)
2nd quintile	25,113 (24.4)	22,675 (23.4)
3rd quintile	21,134 (20.5)	19,506 (20.2)
4th quintile	17,029 (16.5)	17,013 (17.6)
Top quintile	5,728 (5.6)	15,060 (15.6)

Table 2: Variability model for risk of cardiovascular disease

Variable	Adjusted hazard ratio	Standard error	p-value	95% confidence interval	
Female gender	0.560	0.036	<0.001	0.522	0.601
Age (per year)	1.129	0.019	<0.001	1.088	1.171
Atrial fibrillation	2.761	0.159	<0.001	2.022	3.772
Family history of CVD	1.027	0.041	0.512	0.949	1.112
Chronic kidney disease	1.013	0.270	0.963	0.597	1.718
Diabetes	1.654	0.111	<0.001	1.332	2.055
Rheumatoid arthritis	1.709	0.144	<0.001	1.289	2.265
Body mass index (per kg/ m²)	1.009	0.004	0.047	1.000	1.018
TC/HDL cholesterol (per mmol/L)	1.315	0.081	0.001	1.121	1.543
Smoking status (reference = Non-smoker)					
Current smoker	4.458	0.222	<0.001	2.883	6.893
Ex-smoker	1.959	0.278	0.015	1.137	3.374
Ethnic group (reference= White/ Unknown)					
Indian	1.391	0.156	0.034	1.024	1.889
Bangladeshi	1.800	0.711	0.409	0.446	7.255
Pakistani	1.101	0.380	0.799	0.523	2.319
Chinese	0.171	1.001	0.078	0.024	1.216
Black African	1.015	0.290	0.959	0.574	1.794
Black Caribbean	0.644	0.303	0.146	0.355	1.166

Other Black	0.922	0.449	0.857	0.382	2.224
Other Asian	1.350	0.231	0.195	0.857	2.124
Other	0.929	0.215	0.731	0.609	1.415
Mixed	0.970	0.045	0.493	0.888	1.059
Index of multiple deprivation (reference = 1st quintile (least deprived))					
2	1.146	0.041	0.001	1.057	1.243
3	1.192	0.043	<0.001	1.096	1.297
4	1.357	0.044	<0.001	1.243	1.480
5	1.401	0.067	<0.001	1.228	1.598
Mean systolic BP over 6					
readings (per mm Hg)	1.029	0.009	0.001	1.011	1.046
Age x smoking status					
Age in current smokers	0.986	0.004	<0.001	0.979	0.993
Age in ex-smokers	0.992	0.004	0.060	0.983	1.000
Age x TC/HDL cholesterol					
ratio	0.997	0.001	0.066	0.995	1.000
Age x Mean systolic BP over 6					
readings	1.000	0.000	0.064	0.999	1.000
Variation independent of					
mean (per mm Hg)	1.021	0.007	0.005	1.006	1.037

Table 3: External model validation statistics

Statistic	Estimate	95% confidence interval		Median across imputed datasets	Interquartile range across imputed datasets	
Reference model						
c-statistic	0.7415	0.7344	0.7485	0.7415	0.7406	0.7424
R ² (%)	31.84	30.37	33.31	31.85	31.63	32.04
E/O						
calibration	0.9384	0.9113	0.9656	0.9385	0.9365	0.9401
Variability model						
c-statistic	0.7419	0.7348	0.7491	0.7420	0.7412	0.7430
R ² (%)	32.04	30.54	33.53	32.03	31.84	32.23
E/O						
calibration	0.9399	0.9126	0.9673	0.9398	0.9376	0.9417
Across both models						
NRI	-0.0002	-0.0071	0.0068	0.0000	-0.0020	0.0015
IDI	0.0010	0.0000	0.0021	0.0010	0.0007	0.0013

Figure 1: Study flowchart

Figure 2: Observed and predicted risk by decile of risk in the validation cohort: reference model

Figure 3: Observed and predicted risk by decile of risk in the validation cohort: variability model