QRISK INFOSHEET.

Abstract

Objectives To develop and validate updated QRISK3 prediction algorithms to estimate the 10 year risk of cardiovascular disease in women and men accounting for potential new risk factors.

Design Prospective open cohort study.

Setting General practices in England providing data for the QResearch database.

Participants 1309 QResearch general practices in England: 981 practices were used to develop the scores and a separate set of 328 practices were used to validate the scores. 7.89 million patients aged 25-84 years were in the derivation cohort and 2.67 million patients in the validation cohort. Patients were free of cardiovascular disease and not prescribed statins at baseline.

Methods Cox proportional hazards models in the derivation cohort to derive separate risk equations in men and women for evaluation at 10 years. Risk factors considered included those already in QRISK2 (age, ethnicity, deprivation, systolic blood pressure, body mass index, total cholesterol: high density lipoprotein cholesterol ratio, smoking, family history of coronary heart disease in a first degree relative aged less than 60 years, type 1 diabetes, type 2 diabetes, treated hypertension, rheumatoid arthritis, atrial fibrillation, chronic kidney disease (stage 4 or 5)) and new risk factors (chronic kidney disease (stage 3, 4, or 5), a measure of systolic blood pressure variability (standard deviation of repeated measures), migraine, corticosteroids, systemic lupus erythematosus (SLE), atypical antipsychotics, severe mental illness, and HIV/AIDs). We also considered erectile dysfunction diagnosis or treatment in men. Measures of calibration and discrimination were determined in the validation cohort for men and women separately and for individual subgroups by age group, ethnicity, and baseline disease status.

Main outcome measures Incident cardiovascular disease recorded on any of the following three linked data sources: general practice, mortality, or hospital admission records.

Results 363 565 incident cases of cardiovascular disease were identified in the derivation cohort during follow-up arising from 50.8 million person years of observation. All new risk factors considered met the model inclusion criteria except for HIV/AIDS, which was not statistically significant. The models had good calibration and high levels of explained variation and discrimination. In women, the algorithm explained 59.6% of the variation in time to diagnosis of cardiovascular disease (R2, with higher values indicating more variation), and the D statistic was 2.48 and Harrell’s C statistic was 0.88 (both measures of discrimination, with higher values indicating better discrimination). The corresponding values for men were 54.8%, 2.26, and 0.86. Overall performance of the updated QRISK3 algorithms was similar to the QRISK2 algorithms.

Conclusion Updated QRISK3 risk prediction models were developed and validated. The inclusion of additional clinical variables in QRISK3 (chronic kidney disease, a measure of systolic blood pressure variability (standard deviation of repeated measures), migraine, corticosteroids, SLE, atypical antipsychotics, severe mental illness, and erectile dysfunction) can help enable doctors to identify those at most risk of heart disease and stroke.

Methods

Study design and data source

Using the QResearch database (version 41) we undertook a cohort study in a large population of primary care patients. We identified an open cohort of patients aged 25-84 years registered with the practices between 1 January 1998 and 31 December 2015. Patients were excluded if they had no postcode related Townsend score, had pre-existing cardiovascular disease, or were using prescribed statins at cohort entry. Patients were censored at the earliest date of the diagnosis of cardiovascular disease, death, deregistration with the practice, last upload of computerised data, or study end date (31 December 2015).

Outcomes

Our outcome was cardiovascular disease, which was defined as a composite outcome of coronary heart disease, ischaemic stroke, or transient ischaemic attack.

The ICD-10 codes used were G45 (transient ischaemic attack and related syndromes), I20 (angina pectoris), I21 (acute myocardial infarction), I22 (subsequent myocardial infarction), I23 (complications after myocardial infarction), I24 (other acute ischaemic heart disease), I25 (chronic ischaemic heart disease), I63 (cerebral infarction), and I64 (stroke not specified as haemorrhage or infarction). The corresponding ICD-9 codes used were 410, 411, 412, 413, 414, 434, and 436.

Derivation and validation of the models

We used multiple imputation with chained equations to replace missing values for body mass index, systolic blood pressure, standard deviation of systolic blood pressure, serum cholesterol, high density lipoprotein cholesterol, and smoking status and used these values in our main analyses. Five imputations were carried out. In the imputation model we included all predictor variables, along with age interaction terms, the Nelson-Aalen estimator of the baseline cumulative hazard, and the outcome indicator.

Cox’s proportional hazards models were used to estimate the coefficients for each risk factor in women and men separately. We used Rubin’s rules to combine the results across the imputed datasets. Fractional polynomials were used to model non-linear risk relations with continuous variables using data from patients with recorded values to derive the fractional polynomial terms. We fitted full models initially. For consistency, we included variables from existing QRISK2 models and then retained additional variables if they had an adjusted hazard ratio of less than 0.90 or greater than 1.10 (for binary variables) and were statistically significant at the 0.01 level.

From the final models we used the regression coefficients for each variable as weights, which we combined with the baseline survivor function evaluated up to 15 years to derive risk equations over a period of 15 years of follow-up. This enabled us to derive risk estimates for each year of follow-up, with a specific focus on 10 year risk estimates. We estimated the baseline survivor function based on zero values of centred continuous variables, with all binary predictor values set to zero.

Validation of the models

We calculated R2 values (explained variation where higher values indicate a greater proportion of variation in time to cardiovascular disease diagnosis is explained by the model), D statistic (a measure of discrimination where higher values indicate better discrimination), and Harrell’s C statistic at 10 years and combined these across datasets using Rubin’s rules. Harrell’s C statistic is a measure of discrimination that is similar to the area under a receiver operating characteristic curve but takes account of the censored nature of the data.

We assessed calibration (comparing the mean predicted risks at 10 years with the observed risk by 10th of predicted risk). The observed risks were obtained using the Kaplan-Meier estimates evaluated at 10 years. We also evaluated performance in each age group (<40, 40-59, ≥60 years), ethnic origin subgroup, and each comorbidity and treatment subgroup. Performance was also evaluated by calculating Harrell’s C statistics in individual general practices and combining the results using meta-analytical techniques for comparison with a previous study of QRISK2.

Reclassification statistics

In line with current NICE guidelines, we classified patients as being at high risk of cardiovascular disease if their 10 year risk was 10% or greater. We compared predicted risks for our final models (QRISK3) with the latest version of QRISK2-2017 to determine the percentage of patients who would be reclassified at this threshold according to each model. Among the reclassified patients we also calculated the observed risks of cardiovascular disease at 10 years using the Kaplan-Meier method.

Variables used in QRISK3

1. Age
2. Sex
3. Ethnic origin (nine categories)
4. Deprivation (as measured by the Townsend score, where higher values indicate higher levels of material deprivation)
5. Systolic blood pressure
6. Body mass index
7. Total cholesterol : high density lipoprotein cholesterol ratio
8. Smoking status (non-smoker, former smoker, light smoker (1-9/day), moderate smoker (10-19/day), or heavy smoker (≥20/day))
9. Family history of coronary heart disease in a first degree relative aged less than 60 years
10. Diabetes (type 1, type 2, or no diabetes)
11. Treated hypertension (diagnosis of hypertension and treatment with at least one antihypertensive drug)
12. Rheumatoid arthritis (diagnosis of rheumatoid arthritis, Felty’s syndrome, Caplan’s syndrome, adult onset Still’s disease, or inflammatory polyarthropathy not otherwise specified)
13. Atrial fibrillation (including atrial fibrillation, atrial flutter, and paroxysmal atrial fibrillation)
14. Chronic kidney disease (stage 3, 4, or 5) and major chronic renal disease (including nephrotic syndrome, chronic glomerulonephritis, chronic pyelonephritis, renal dialysis, and renal transplant)
15. Measure of systolic blood pressure variability (standard deviation of repeated measures)
16. Diagnosis of migraine
17. Corticosteroid use (British National Formulary (BNF) chapter 6.3.2 including oral or parenteral prednisolone, betamethasone, cortisone, depo-medrone, dexamethasone, deflazacort, efcortesol, hydrocortisone, methylprednisolone, or triamcinolone)
18. Systemic lupus erythematosus (including diagnosis of SLE, disseminated lupus erythematosus, or Libman-Sacks disease)
19. Second generation “atypical” antipsychotic use (including amisulpride, aripiprazole, clozapine, lurasidone, olanzapine, paliperidone, quetiapine, risperidone, sertindole, or zotepine)
20. Diagnosis of severe mental illness (including psychosis, schizophrenia, or bipolar affective disease)
21. Diagnosis of erectile dysfunction or treatment for erectile dysfunction (BNF chapter 7.4.5 including alprostadil, phosphodiesterase type 5 inhibitors, papaverine, or phentolamine)

Results

In the derivation cohort, we identified 363 565 incident cases of cardiovascular disease arising from 50.8 million person years of observation. The incidence of cardiovascular disease increased steeply by age group and values were higher in men than women for all age groups. Of the 363 565 incident events, 78 327 (21.5%) were myocardial infarction, 152 141 (41.8%) were angina, 49 504 (13.6%) were transient ischaemic attack, and 83 593 (23.0%) were ischaemic strokes. Overall, 92 936 (25.6% of all 363 565 events) were only recorded on the general practice record, with the most common condition being transient ischaemic attack (27 227 events).

Predictor variables

For the new variables of interest in model B, migraine was associated with a 36% increased risk of cardiovascular disease for women and a 29% increased risk for men, corticosteroids were associated with an 82% increased risk for women and 58% increased risk for men, SLE was associated with a 115% increased risk for women and a 55% increased risk for men, atypical antipsychotics were associated with a 29% increased risk for women and a 15% increased risk for men, severe mental illness was associated with a 14% increased risk for women and a 13% increased risk for men. Erectile dysfunction was associated with a 25% increased risk. Where there were age interactions these values relate to risks evaluated at the mean ages.

For the new variables, there were statistically significant interactions between age and migraine as well as age and corticosteroid use in both sexes. In women, there was also a statistically significant interaction between age and SLE. In men, there was also a statistically significant interaction between age and erectile dysfunction. For each of these interactions, hazard ratios for the predictors were higher at younger ages compared with older ages, except for erectile dysfunction in men, where hazard ratios were highest for men aged around age 45 and then declined gradually with increasing age.

For model C, the standard deviation of systolic blood pressure values was included in the model in addition to the single most recent systolic blood pressure value. Overall a 10 unit increase in the standard deviation of systolic blood pressure was associated with an 8% increased risk of cardiovascular disease in women and an 11% increased risk in men.

Validation

Discrimination

For model B in women, the algorithm explained 59.5% of the variation in time to diagnosis of cardiovascular disease (R2), the D statistic was 2.48, and the Harrell’s C statistic was 0.88. The corresponding values for men were 54.8%, 2.26, and 0.86. Measures of performance were similar for all three models.

The highest performance values by ethnic origin were in Chinese women (R2=64.7%; D=2.77; Harrell’s C=0.91) and the lowest values were in Caribbean women (R2=51.6%; D=2.11; Harrell’s C=0.85). Performance values were highest in the youngest age group (25-39 years) and lowest in the oldest age group (60-84 years). For the subgroup of women with type 1 diabetes the R2 was 47.3%, D statistic was 1.94, and Harrell’s C statistic was 0.82. The corresponding values for men with type 1 diabetes were 45.6%, 1.87, and 0.80. For the subgroup of women with type 2 diabetes the R2 was 25.2%, D statistic was 1.19, and Harrell’s C statistic was 0.70. The corresponding values for men with type 2 diabetes were 22.9%, 1.12, and 0.70.

Calibration

In women, the mean 10 year predicted risk was 4.7% and the observed 10 year risk was 5.8% (95% confidence interval 5.8% to 5.9%). In men, the mean 10 year predicted risk was 6.4% and the observed 10 year risk was 7.5% (7.5% to 7.6%). There was close correspondence between the mean predicted risks and the observed risks within each model 10th overall and in each age group in women and men indicating that the algorithms were well calibrated. The exception was in those aged 25-39 where mean predicted risks were slightly higher than observed risks.

Discussion

We have developed and validated updated algorithms (QRISK3) to predict 10 year risk of cardiovascular disease in women and men aged 25-84 years. The algorithms incorporate established predictor variables from QRISK2 as well as new variables associated with increased risk of cardiovascular disease. These include an expanded definition of chronic kidney disease to include chronic kidney disease stage 3, migraine, corticosteroid use, systemic lupus erythematosus (SLE), atypical antipsychotic use, severe mental illness, erectile dysfunction, and a measure of blood pressure variability (standard deviation of repeated values). Although in population terms the overall performance of all three models is similar, for those who have one or more of the conditions included in the newer models, having the additional risk taken into account could result in the difference between taking or not taking risk reducing treatment. The increased complexity is unlikely to affect the take-up of the new models as they are designed to be calculated automatically from the electronic patient record.

**IMPACT OF DIFFERENT VARIABLES ON RISK FOR QRISK3 MODELS.**

The Qrisk3 table presents the effects of different covariates on underlying outcome risk via hazard ratios. Generally, men are at higher risk than women in the QRISK3 model. We present the results of two tables: for women and for men.

The tables present hazard ratios for cardiovascular disease in women and men. For women, the Townsend score shows a hazard ratio of 1.47 (1.45 to 1.50) per 5 unit increase. Regarding ethnic origin compared to White, Indian women have a ratio of 1.32 (1.26 to 1.39), Pakistani 1.76 (1.66 to 1.86), Bangladeshi 1.34 (1.24 to 1.45), Other Asian 1.08 (0.992 to 1.17), Black Caribbean 0.843 (0.797 to 0.891), Black African 0.675 (0.618 to 0.737), Chinese 0.722 (0.622 to 0.837), and Other 0.843 (0.791 to 0.897). For smoking status compared to non-smokers, former smokers have a ratio of 1.14 (1.11 to 1.18), light smokers 1.75 (1.70 to 1.81), moderate smokers 1.95 (1.88 to 2.02), and heavy smokers 2.34 (2.25 to 2.43). Medical characteristics show family history of coronary heart disease at 1.58 (1.54 to 1.61), Type 1 diabetes 5.62 (5.08 to 6.22), Type 2 diabetes 2.91 (2.72 to 3.11), treated hypertension 1.66 (1.60 to 1.73), rheumatoid arthritis 1.24 (1.20 to 1.27), atrial fibrillation 4.92 (4.20 to 5.75), chronic kidney disease (stage 3, 4, or 5) 1.92 (1.70 to 2.17), migraine 1.35 (1.30 to 1.40), corticosteroid use 1.81 (1.74 to 1.89), and systemic lupus erythematosus 2.14 (1.78 to 2.56).

For men, the Townsend score shows a hazard ratio of 1.18 (1.17 to 1.20) per 5 unit increase. Ethnic origin ratios for men compared to White are: Indian 1.32 (1.27 to 1.37), Pakistani 1.61 (1.53 to 1.68), Bangladeshi 1.70 (1.61 to 1.79), Other Asian 1.04 (0.970 to 1.11), Black Caribbean 0.699 (0.662 to 0.738), Black African 0.670 (0.623 to 0.721), Chinese 0.660 (0.582 to 0.749), and Other 0.769 (0.728 to 0.812). For smoking status in men, former smokers have a ratio of 1.21 (1.18 to 1.24), light smokers 1.74 (1.70 to 1.78), moderate smokers 1.89 (1.84 to 1.94), and heavy smokers 2.20 (2.14 to 2.27). Medical characteristics for men show family history of coronary heart disease at 1.72 (1.69 to 1.75), Type 1 diabetes 3.44 (3.17 to 3.73), Type 2 diabetes 2.36 (2.23 to 2.50), treated hypertension 1.68 (1.61 to 1.74), rheumatoid arthritis 1.23 (1.19 to 1.28), atrial fibrillation 2.42 (2.14 to 2.73), chronic kidney disease (stage 3, 4, or 5) 2.05 (1.83 to 2.29), migraine 1.29 (1.24 to 1.34), corticosteroid use 1.58 (1.5 to 1.66), systemic lupus erythematosus 1.55 (1.15 to 2.10), and erectile dysfunction or treatment 1.25 (1.18 to 1.33).