

Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2

Julia Hippisley-Cox,¹ Carol Coupland,¹ Yana Vinogradova,¹ John Robson,² Rubin Minhas,³ Aziz Sheikh,⁴ Peter Brindle⁵

EDITORIAL by Christiaens

¹Division of Primary Care, Tower Building, University Park, Nottingham NG2 7RD

²Centre for Health Sciences, Queen Mary's School of Medicine and Dentistry, London E1 2AT

³Medway Primary Care Trust, Unit 2, Gillingham, Kent ME7 0NJ

⁴Division of Community Health Sciences: GP Section, University of Edinburgh, Edinburgh EH8 9DX

⁵Avon Primary Care Research Collaborative, Bristol Primary Care Trust, Bristol BS2 8EE

Correspondence to: J Hippisley-Cox
julia.hippisley-cox@nottingham.ac.uk

BMJ 2008;336:1475-82

doi:10.1136/bmj.39609.449676.25

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: BMJ 2008, doi:10.1136/bmj.39609.449676.25

ABSTRACT

Objective To develop and validate version two of the QRISK cardiovascular disease risk algorithm (QRISK2) to provide accurate estimates of cardiovascular risk in patients from different ethnic groups in England and Wales and to compare its performance with the modified version of Framingham score recommended by the National Institute for Health and Clinical Excellence (NICE).

Design Prospective open cohort study with routinely collected data from general practice, 1 January 1993 to 31 March 2008.

Setting 531 practices in England and Wales contributing to the national QRESEARCH database.

Participants 2.3 million patients aged 35-74 (over 16 million person years) with 140 000 cardiovascular events. Overall population (derivation and validation cohorts) comprised 2.22 million people who were white or whose ethnic group was not recorded, 22 013 south Asian, 11 595 black African, 10 402 black Caribbean, and 19 792 from Chinese or other Asian or other ethnic groups.

Main outcome measures First (incident) diagnosis of cardiovascular disease (coronary heart disease, stroke, and transient ischaemic attack) recorded in general practice records or linked Office for National Statistics death certificates. Risk factors included self assigned ethnicity, age, sex, smoking status, systolic blood pressure, ratio of total serum cholesterol:high density lipoprotein cholesterol, body mass index, family history of coronary heart disease in first degree relative under 60 years, Townsend deprivation score, treated hypertension, type 2 diabetes, renal disease, atrial fibrillation, and rheumatoid arthritis.

Results The validation statistics indicated that QRISK2 had improved discrimination and calibration compared with the modified Framingham score. The QRISK2 algorithm explained 43% of the variation in women and 38% in men compared with 39% and 35%, respectively, by the modified Framingham score. Of the 112 156 patients classified as high risk (that is, $\geq 20\%$ risk over 10 years) by the modified Framingham score, 46 094 (41.1%) would be reclassified at low risk with QRISK2. The 10 year observed risk among these reclassified patients was 16.6% (95% confidence interval 16.1% to 17.0%)—that is, below the 20% treatment threshold. Of the 78 024

patients classified at high risk on QRISK2, 11 962 (15.3%) would be reclassified at low risk by the modified Framingham score. The 10 year observed risk among these patients was 23.3% (22.2% to 24.4%)—that is, above the 20% threshold. In the validation cohort, the annual incidence rate of cardiovascular events among those with a QRISK2 score of $\geq 20\%$ was 30.6 per 1000 person years (29.8 to 31.5) for women and 32.5 per 1000 person years (31.9 to 33.1) for men. The corresponding figures for the modified Framingham equation were 25.7 per 1000 person years (25.0 to 26.3) for women and 26.4 (26.0 to 26.8) for men). At the 20% threshold, the population identified by QRISK2 was at higher risk of a CV event than the population identified by the Framingham score.

Conclusions Incorporating ethnicity, deprivation, and other clinical conditions into the QRISK2 algorithm for risk of cardiovascular disease improves the accuracy of identification of those at high risk in a nationally representative population. At the 20% threshold, QRISK2 is likely to be a more efficient and equitable tool for treatment decisions for the primary prevention of cardiovascular disease. As the validation was performed in a similar population to the population from which the algorithm was derived, it potentially has a "home advantage." Further validation in other populations is therefore advised.

INTRODUCTION

Recent advances in the development of models to assess risk of cardiovascular disease now take account of the increased risk associated with social deprivation in the UK.^{1,2} Rates of cardiovascular disease, however, vary considerably between ethnic groups, which might reflect increased susceptibility and differential exposure to risk factors. NICE recommended multiplying the results of a modified version of the US Framingham score ("modified Framingham") by a correction factor of 1.4 for south Asian men in the UK.³ This does not reflect the heterogeneity in risk of cardiovascular disease between south Asian populations, the increased risk in women, confounding by deprivation,⁴ and the possibility of double counting through adjustments for both ethnicity and family history.

We built on our previous risk prediction algorithm (QRISK1)¹ to develop a revised algorithm that incorporates self assigned ethnicity as well as a range of other potentially relevant conditions associated with cardiovascular risk such as type 2 diabetes, treated hypertension, rheumatoid arthritis, renal disease, and atrial fibrillation (QRISK2). By including an increased range of potential risk factors, we hypothesised that we would be better able to personalise risk to the individual patient.

METHODS

Study design and data source

We conducted a prospective cohort study in a large UK primary care population using a similar method to our original analysis.¹ We used version 19 of the QRESEARCH database (www.qresearch.org). This is a large validated primary care electronic database containing the health records of 11 million patients registered from 551 general practices using the Egton Medical Information System (EMIS) computer system.¹ Practices and patients on the database are nationally representative and similar to those on other primary care databases that use other clinical software systems.⁵

The QRESEARCH database now contains information on the cause of death as recorded on the patient's Office for National Statistics (ONS) death certificate. A recorded cause of death is now linked for over 97% of patients on the QRESEARCH database who have died.

Practice selection—We included all QRESEARCH practices in England and Wales once they had been using their current EMIS system for at least a year, randomly allocating two thirds of practices to the derivation dataset with one third to the validation dataset.

Cohort selection—We identified an open cohort of patients aged 35-74 at the study entry date, drawn from patients registered with eligible practices from 1

January 1993 to 31 March 2008. We excluded patients with a prior recorded diagnosis of cardiovascular or cerebrovascular disease, temporary residents, patients with interrupted periods of registration with the practice, and those who did not have a valid Townsend deprivation score. We also excluded patients who were taking statins at baseline.

Coding of ethnicity—We used Read codes for self assigned ethnicity. The codes were grouped into the NHS standard 16+1 categories.⁶ These categories were then further grouped into the final nine reporting groups to ensure sufficient numbers of events to enable a meaningful analysis. See bmj.com.

Cardiovascular disease outcomes

The primary outcome measure was the first recorded diagnosis of cardiovascular disease recorded on the general practice clinical computer system or their linked ONS death certificate during the study period. We included coronary heart disease (angina and myocardial infarction), stroke, or transient ischaemic attacks in the term cardiovascular disease but not peripheral vascular disease.

The Read codes used for case identification on the computer record were nationally agreed ones used in the quality and outcomes framework for general practice for coronary heart disease and cerebrovascular disease.

Risk factors for cardiovascular disease

Variables included in our analysis are shown in the box.

Model derivation and development

We calculated crude incidence rates of cardiovascular disease according to age, ethnic group, and deprivation in fifths. We directly age standardised the incidence rates by ethnic group and deprivation using the age distribution in five year bands of the entire derivation cohort as the standard population. We also age standardised the means of continuous variables and proportions with risk factors by ethnic group using the same method.

We used Cox proportional hazards models in the derivation dataset to estimate the coefficients and hazard ratios associated with each potential risk factor for the first ever recorded diagnosis of cardiovascular disease for men and women separately. We compared models using the Bayesian information criteria (BIC). We tested for interactions between each variable and age and between diabetes and deprivation and included significant interactions in the final model. Our main analyses used multiple imputation to replace missing values for systolic blood pressure, cholesterol/HDL ratio, smoking status, and body mass index. Our final model was fitted based on multiple imputed datasets. See bmj.com.

We took the log of the hazard ratio for each variable from the final model and used these as weights for the new cardiovascular disease risk equations. We

Included variables

- Self assigned ethnicity (white/not recorded, Indian, Pakistani, Bangladeshi, other Asian, black African, black Caribbean, Chinese, other including mixed)
- Age (years)
- Sex (males v females)
- Smoking status (current smoker, non-smoker (including ex-smoker))
- Systolic blood pressure⁷ (continuous)
- Ratio of total serum cholesterol/high density lipoprotein cholesterol⁷ (continuous)
- Body mass index (BMI)¹ (continuous)
- Family history of coronary heart disease in first degree relative under 60 years¹ (yes/no)
- Townsend deprivation score¹ (output area level 2001 census data evaluated as a continuous variable)
- Treated hypertension¹ (diagnosis of hypertension and at least one current prescription of at least one antihypertensive agent)
- Rheumatoid arthritis⁸ (yes/no)
- Chronic renal disease⁹ (yes/no)
- Type 2 diabetes⁷ (yes/no)
- Atrial fibrillation^{10,11} (yes/no)

combined these weights with the baseline survivor function centred on the means of continuous risk factors to derive a risk equation for 10 years' follow-up.

Validation of new equation

We tested the performance of the new model (QRISK2) in the validation dataset and compared it against both the original model (QRISK1) and the modified Framingham equation recommended by NICE.¹² We calculated the 10 year estimated risk of cardiovascular disease for each patient in the validation dataset using multiple imputation to replace missing values as in the derivation dataset.

We calculated the mean predicted and observed cardiovascular disease risk at 10 years¹ and compared these by 10th of predicted risk for each score. The observed risk at 10 years was obtained by using the 10 year Kaplan-Meier estimate. We calculated the Brier score (a measure of goodness of fit where lower values indicate better accuracy), D statistic (a measure of discrimination where higher values indicate better discrimination), and an R² statistic. The R² statistic is a measure of explained variation where higher values indicate more explained variation. We also calculated the area under the receiver operator curve (ROC), where higher values indicate better discrimination.

We calculated the proportion of patients in the validation sample with an estimated 10 year risk of cardiovascular disease of 20% or more by age, sex, ethnicity, and deprivation according to the QRISK2 algorithm compared with the modified Framingham score.

RESULTS

Derivation and validation datasets

Practices and patients—Overall, 531 UK practices met our inclusion criteria, of which 355 were randomly assigned to the derivation dataset and 176 to the validation dataset. We excluded 20 practices. We studied 2.29 million patients with over 16 million person years and 140 115 cardiovascular events. There were 1 591 209 patients in the derivation cohort, of whom 55 626 had cardiovascular disease before the start of the study leaving 1 535 583 patients (773 291

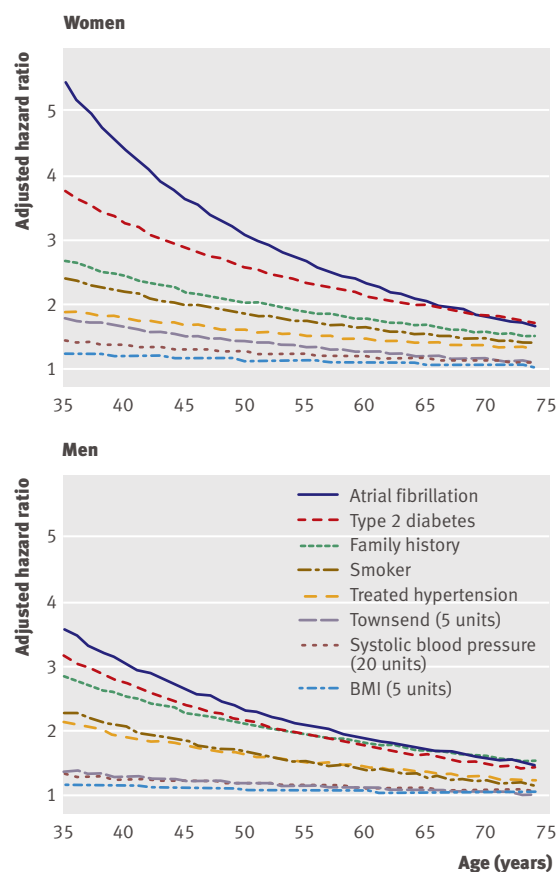


Fig 1 | Impact of age on hazard ratios for cardiovascular disease risk factors using the QRISK2 model

women, 50.4%) aged 35-74 and free of cardiovascular disease. See bmj.com.

Baseline characteristics of derivation and validation cohort—Ethnicity was recorded in 209 214 (27.1%) women and 181 110 (23.8%) men. Among patients with ethnicity recorded 89.3% were from a white ethnic group. The mean follow-up was 7.3 years for women and 6.9 for men. The baseline characteristics of the validation cohort were similar to those for the derivation cohort.

Incidence of cardiovascular disease—There were 96 709 incident cases of cardiovascular disease (41 042 in women) during the study period from 10.9 million person years of observation. Of all events, 7.4% in women and 7.8% in men were identified with the ONS linked death data (that is, were not identified by using the general practice data alone). The crude incidence rate for cardiovascular disease was slightly higher than in our original study with a rate of 7.3 per 1000 person years for women and 10.5 per 1000 person years for men. In the validation dataset there were 750 232 eligible patients aged 35 to 74, and, of these, 50.1% were women and the incidence rates were similar to the derivation dataset. The age standardised rates for the white reference group were 10.5 per 1000 person years (95% confidence interval 10.4 to 10.6) for men and 7.3 per 1000 person years (7.2 to 7.3) for women. The highest age standardised rates were among south Asian

Table 1 | Validation statistics for new QRISK2 model compared with modified NICE equation in validation cohort. Figures are means (95% confidence intervals)

| | QRISK2 model | QRISK1 model | Modified Framingham equation |
|----------------|------------------------|------------------------|------------------------------|
| Women | | | |
| R ² | 43.47 (42.78 to 44.16) | 42.94 (42.23 to 43.66) | 38.87 (38.12 to 39.62) |
| D statistic | 1.795 (1.769 to 1.820) | 1.776 (1.750 to 1.801) | 1.632 (1.606 to 1.658) |
| ROC statistic | 0.817 (0.814 to 0.820) | 0.814 (0.811 to 0.817) | 0.800 (0.797 to 0.803) |
| Brier score | 0.086 (0.083 to 0.089) | 0.081 (0.078 to 0.084) | 0.093 (0.090 to 0.096) |
| Men | | | |
| R ² | 38.38 (37.75 to 39.01) | 37.63 (36.99 to 38.27) | 34.78 (34.12 to 35.45) |
| D statistic | 1.615 (1.594 to 1.637) | 1.590 (1.568 to 1.612) | 1.495 (1.473 to 1.517) |
| ROC statistic | 0.792 (0.789 to 0.794) | 0.788 (0.786 to 0.791) | 0.779 (0.776 to 0.782) |
| Brier score | 0.136 (0.134 to 0.139) | 0.128 (0.125 to 0.131) | 0.177 (0.174 to 0.180) |

groups. Age standardised rates were also high for Indian and Pakistani men and women compared with the white reference group. They were also higher for black Caribbean women and men from the other Asian group. In contrast, black African, Chinese, and black Caribbean men tended to have lower rates, as did black African women. See bmj.com.

Characteristics of events

Overall, 30.8% of events were stroke or transient ischaemic attacks, but this varied between ethnic groups. For example in the derivation dataset, 48.9% of first events among black Caribbean men and 36.4% among black African men were stroke or transient ischaemic attacks; the corresponding figures for women were 33.5% and 24.2%.

Prevalence of risk factors by ethnicity

There was substantial heterogeneity across the ethnic groups in risk factors for cardiovascular disease and this also differed between men and women within an ethnic group. The notable results include differences in the age standardised prevalence of smoking among men of Bangladeshi (53.2%, 50.2% to 56.2%), Caribbean

(40.6%, 38.9% to 42.4%), Pakistani (32.9%, 30.8% to 35.1%), white/not recorded (32.2%, 32.1% to 32.3%), Chinese (28.0%, 24.6% to 31.4%), Indian (23.7%, 22.3% to 25.1%), and black African (16.6%, 15.1% to 18.2%) origin. Current smoking rates were all lower for women in each ethnic group compared with men but varied widely between women from different groups.

There were also substantial differences in the age standardised prevalence of type 2 diabetes between ethnic groups with highest rates among Bangladeshis (14.4% women, 16.8% men), Pakistanis (14.2% women, 12.0% men), and Indians (11.7% women, 13.3% men) and lowest among the white reference group (1.5% women, 2.1% men).

Treated hypertension was highest among Caribbean and black African men and women. Recorded family history of coronary heart disease in a first degree relative was highest among Indian men and women and lowest among black African men and women.

Model development

In the Cox regression analysis for the QRISK2 model, we used a log transformation for age but otherwise fitted variables as linear terms. See bmj.com for table of

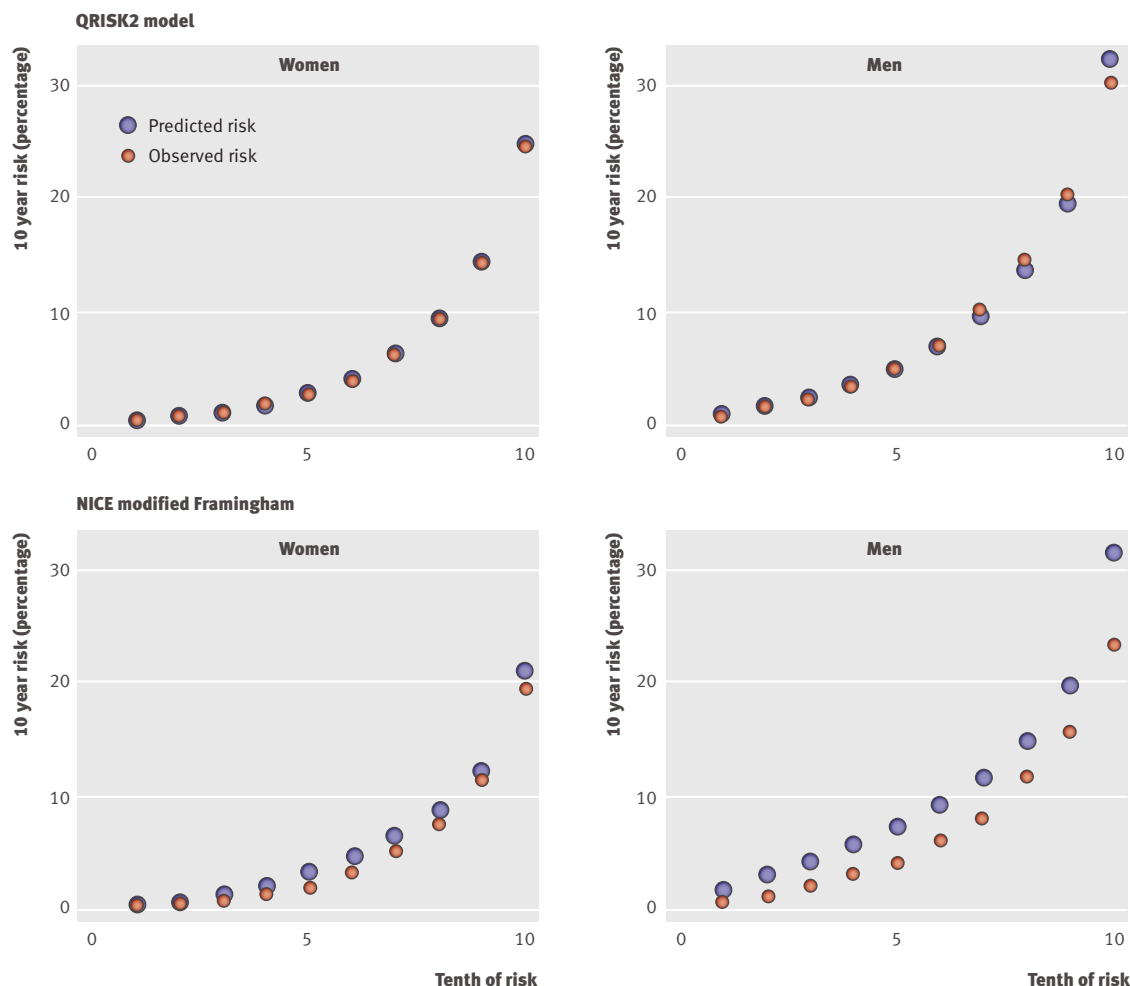


Fig 2 | Predicted and observed risk by 10th of predicted risk for QRISK2 model and NICE modification of Framingham score in the validation dataset

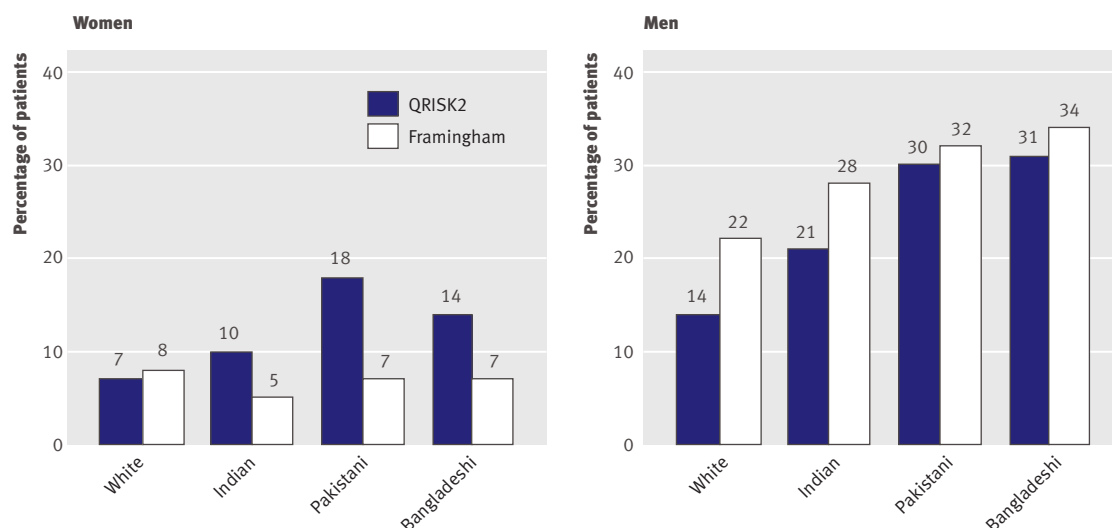


Fig 3 | Percentage of white and south Asian patients at high risk in validation dataset with QRISK2 and Framingham score in QRESEARCH database (in modified Framingham score, inflation factor of 1.4 is applied to south Asian men but not south Asian women)

adjusted hazard ratios. Figure 1 shows the impact of age on hazard ratios.

Calibration and discrimination of QRISK2

The QRISK2 model was marginally superior to the original QRISK1 equation and both models were superior to the modified Framingham score for the D statistic, ROC statistic, and the R^2 value—for both men and women (table 1).

Figure 2 compares predicted and observed risks of a cardiovascular disease event at 10 years across each 10th of predicted risk (first 10th representing the lowest risk). The QRISK2 model is better calibrated than the modified Framingham score.

Predictions with age, sex, deprivation, and ethnicity

Overall, the QRISK2 model would predict 10.6% of patients as high risk (risk of $\geq 20\%$ over 10 years) compared with 14.9% for the modified Framingham score. See bmj.com. Figure 3 shows the proportion of patients estimated to be at high risk with QRISK2 and the Framingham score within each ethnic group.

Reclassification statistics

Of the 112 156 patients classified as high risk (risk of $\geq 20\%$ over 10 years) with the Framingham score, 46 094 (41.1%) would be reclassified at low risk with QRISK2. The 10 year observed risk among these reclassified patients was 16.6% (16.1% to 17.0%)—that is, below the 20% threshold for high risk.

Of the 78 024 patients classified at high risk with QRISK2, 11 962 (15.3%) would be reclassified as low risk with the Framingham score. The 10 year observed risk among these patients predicted to be at high risk with QRISK2 was 23.3% (22.2% to 24.4%)—that is, above the 20% threshold for high risk.

The annual incidence rate of cardiovascular events among those with a QRISK2 score of $\geq 20\%$ was 30.6

per 1000 person years (95% confidence interval 29.8 to 31.5) for women and 32.5 per 1000 person years (31.9 to 33.1) for men. Both these figures are higher than the annual incidence rate for patients identified as high risk with the modified Framingham score. The annual incidence rate for these patients was 25.7 per 1000 person years (25.0 to 26.3) for women with 26.4 (26.0 to 26.8) for men. In other words, at the 20% threshold, the population identified by QRISK2 was at higher risk of a CV event than the population identified by the NICE modified Framingham algorithm.

Clinical examples

Table 2 shows some clinical examples for patients from different ethnic groups who would be reclassified with QRISK2 compared with the modified Framingham score.

DISCUSSION

We developed and validated a cardiovascular risk algorithm that simultaneously takes account of ethnicity and deprivation and provides an individualised estimate of cardiovascular risk. It also extends and improves on our original equation for cardiovascular risk¹ by incorporating important additional clinical conditions. This information should be considered in the context of specific treatment guidelines. It also allows better quantification of risk of cardiovascular disease for patients with type 2 diabetes, which is especially prevalent among south Asian patients. Although current guidelines might indicate statins for people with diabetes, knowledge of cardiovascular risk can identify patients at particularly low risk for whom a statin might not be needed.

Strengths and limitations

The strengths and limitations of using this approach and the QRESEARCH database to develop and

validate a new risk prediction algorithm have been discussed previously.¹⁵ We included more sophisticated modelling of the effect of age on risk factors, which results in greater weighting of some risk factors in younger patients. The inclusion of patients with type 2 diabetes will have tended to increase the overall level of risk in the study population and this will also have tended to increase the risk for an individual.

We updated the analysis to include data until March 2008, increasing the number of patients with at least 10 years of follow-up data. We have furthermore included the linked cause of death as recorded by the Office for National Statistics (ONS).

We used self assigned ethnicity as reported by the patient to their general practice; this has advantages over analyses where ethnicity is assigned by an informant rather than the patient or is imputed geographically or is related to country of birth. We also disaggregated the south Asian groups and reported on them separately. Misclassification would most affect the reference category of “white or not recorded,” but because of the mix of the populations of England and Wales less than 10% of such patients were probably from a non-white ethnic group. This misclassification would therefore, if anything, tend to underestimate the relative effect of ethnicity on cardiovascular risk.

Just fewer than 3% of our total sample were classified as belonging to a minority ethnic group compared with the national proportion in this age group of 6.6% (based on projections for 2006). However, national estimates are for 2006 and migration patterns and population demographics have probably changed over the 15 year

period of our study. None the less, the lower percentage of patients from minority groups raises concerns about the possible under-representativeness of practices from ethnically diverse inner city areas or misclassification error, or both. See bmj.com for discussion.

We have assumed that the absence of a recorded diagnosis of diabetes (or family history, for example) is equivalent to the person not having that factor. As recording of risk factors becomes more complete over time, then better estimates of the relevant hazard ratios will be possible.

We have calculated 95% confidence intervals around the QRISK2 scores to give a better idea of precision. We have improved on the method for validation by using multiple imputation for missing values in the validation set.¹⁵ One important limitation is that while we have validated the results in a physically discrete group of practices, these practices all use the same EMIS software which might reduce the generalisability. It is important that QRISK2 is validated by another team on external populations and an international version of QRISK2 is being developed. We are also working with another primary care database (THIN, “The Health Improvement Network”) so that this can be used as a data source for further validation. Ethnicity recording could be improved on primary care databases by linkage of individual level data on self assigned ethnicity from the 2001 census.

Comparisons with the modified Framingham score
With QRISK2, the improvement in discrimination and calibration compared with the modified Framingham

Table 2 | Clinical examples for patients who would be reclassified with QRISK2 instead of NICE modified Framingham equation

| Age (years) | Ethnic group | Family history | Systolic blood pressure | BMI | Cholesterol/HDL ratio | Smoker | Treated hypertension | Type 2 diabetes* | Chronic kidney disease | Townsend score† | Framingham score 10 year risk (%) | QRISK2 10 year risk (%) (95% CI) |
|--------------|---------------|----------------|-------------------------|------|-----------------------|--------|----------------------|------------------|------------------------|-----------------|-----------------------------------|----------------------------------|
| Men | | | | | | | | | | | | |
| 65 | Indian | Yes | 100 | 24.7 | 3.3 | No | No | No | No | 5 | 17 | 31.3 (30.9 to 31.7) |
| 54 | Bangladeshi | No | 142 | 27.0 | 4.2 | No | Yes | No | No | 10 | 17 | 23.5 (22.8 to 24.1) |
| 54 | Black African | No | 150 | 21.0 | 7.3 | No | No | No | No | 4 | 23 | 9.0 (7.7 to 10.3) |
| 55 | Indian | No | 156 | 27.0 | 4.7 | No | No | No | No | -4 | 24 | 12.7 (12.2 to 13.2) |
| 65 | Caribbean | No | 146 | 29.1 | 5.4 | No | No | No | No | 4 | 26 | 14.8 (14.2 to 15.5) |
| 42 | White | Yes | 132 | 36.0 | 5.3 | Yes | Yes | No | No | 11 | 17 | 35.2 (34.9 to 35.5) |
| Women | | | | | | | | | | | | |
| 64 | Indian | No | 130 | 23.1 | 5.3 | No | Yes | No | No | 5 | 12 | 24.7 (24.4 to 25.0) |
| 60 | Bangladeshi | No | 132 | 36.0 | 4.3 | No | Yes | No | No | 11 | 9 | 21.1 (20.6 to 21.6) |
| 48 | Pakistani | Yes | 140 | 33.2 | 4.5 | No | Yes | No | No | 8 | 9 | 26.1 (25.7 to 26.4) |
| 58 | White | No | 154 | 34.0 | 3.4 | Yes | Yes | No | No | 10 | 16 | 21.4 (21.3 to 21.5) |

BMI=body mass index; HDL=high density lipoprotein cholesterol.

*NICE lipid modification guideline does not include diabetes so this is for illustrative purposes only.

†Interval score ranges between -6 (most affluent) and 11 (most deprived).

WHAT IS ALREADY KNOWN ON THIS TOPIC

A 10 year cardiovascular disease risk threshold of 20% is recommended for intervention with statins for the primary prevention of cardiovascular disease

Current algorithms for risk of cardiovascular disease do not adequately account for the combined effect of socioeconomic status and ethnicity, leading to an underestimate of risk in high risk populations that might potentially exacerbate existing health inequalities

WHAT THIS STUDY ADDS

Compared with a white reference population, there is a substantially increased risk of cardiovascular disease in south Asian men and women that is independent of social deprivation, diabetes, and family history

The results of the calibration and discrimination statistics for QRISK2 were significantly better than those for the modified Framingham score in the validation sample

At the 10 year risk threshold of 20%, the population identified by QRISK2 was at higher risk of a CV event than the population identified by the modified algorithm

score remains significant, although this is probably partly because the modelling was undertaken on a more contemporaneous population from England and Wales and we used a more sophisticated approach for modelling and included additional variables. We have not compared QRISK2 with the most recently published Framingham score as this uses a much broader definition of cardiovascular disease that is less relevant to UK guidelines.¹³ QRISK2 seems to improve on the Framingham score based Ethrisk,¹⁴ perhaps because of its greater precision, larger sample, and prospective study design.

We compared QRISK2 with the modified Framingham risk score recently recommended by NICE. The modified score involves summing risks from two risk equations for coronary heart disease and stroke, which is mathematically incorrect because these are not independent outcomes and therefore will give an invalid result. The inflation factors for 1.4 for south Asian men and 1.5 for those with a family history of coronary heart disease, which have been developed by consensus rather than a mathematical model based on individual patient data, might also have accounted for some of the overprediction, although the overprediction was still present on our previous analysis where the inflation factors for the Framingham score had not been applied.¹⁵

Comparisons with the literature

We found substantial heterogeneity between risk factors within south Asian populations and our prevalence figures for risk factors are comparable with the literature,^{4,14} which increases the face validity of our findings. Our findings also confirm Nazroo's observations¹⁵ and the findings of the Whitehall II study¹⁶ of the independent effects of both ethnicity and deprivation. The magnitude of the increased

cardiovascular risk among south Asians compared with white patients seems to be higher than the 40% previously thought in the absence of prospective incidence data.^{17,18}

There were also differences in the proportion of events that were stroke or transient ischaemic attacks rather than coronary heart disease among different ethnic groups, which is consistent with the literature and deserves further study.

Clinical implementation

The mapping of postcode to deprivation score will be made available, together with the supporting reference tables and QRISK2 algorithm itself. QRISK2 can then be integrated within clinical management systems so that it can be used on an ongoing basis to generate an estimated score based on existing data. The plan for electronic health records to eventually incorporate computerised decision support tools will allow disease risk algorithms such as QRISK2 to be largely automatically populated with routine electronically coded data as is already possible in primary care in the UK.

QRISK2 provides a mechanism for estimating absolute risk among individuals. Use of this information, however, should be tightly coupled with suitable guidelines. There are some patients in whom a QRISK2 score should not be calculated, including those with pre-existing cardiovascular disease (who we excluded from this study).

Clinical impacts and health inequalities

The clinical relevance, superior performance, and equitable assignment of QRISK2 make it an appropriate tool to assist in the delivery of public health programmes that recognise the broader determinants of cardiovascular health, such as ethnicity and deprivation. This has particular relevance to equity of delivery of health care to the UK's south Asian communities and might help to reduce widening health inequalities.

We acknowledge the contribution of David Stables (EMIS) and EMIS practices contributing to the QRESEARCH database. In particular we acknowledge his contribution in linking the ONS death certificate data to individual records held within EMIS clinical systems so that it could be extracted on to the QResearch database and used for this project. We thank Aneez Esmail (University of Manchester), Ruthie Birger and Chris Millett (Imperial College London), and Nadeem Qureshi (University of Nottingham) for ethnicity coding.

Contributors: See bmj.com.

Funding: No external funding. The authors were funded as part of their clinical or academic positions and meeting expenses were met by the University of Nottingham.

Competing interests: JR chaired and PB and RM were members of the NICE guideline development group on cardiovascular risk assessment. JHC is codirector of QRESEARCH—a not for profit organisation that is a joint partnership between the University of Nottingham and EMIS. EMIS is the leading commercial supplier of IT systems for 56% of general practices in England and Wales and it is likely to implement QRISK2 into its clinical management system. EMIS is likely to also distribute the software package for those using it for academic research or other organisations interested in implementing QRISK2 into practice or (www.qresearch.org/Public/qriskInformationforClinicians.aspx). RM is a 2008 Harkness Fellow in healthcare policy and practice and is the chair of the cardiovascular working group of the South Asian Health Foundation (SAHF), which receives unrestricted funding from the Department of Health and BHF and

unrestricted grants from the pharmaceutical industry. AS chairs the equality and diversity forum of the National Clinical Assessment Service. AS is PI on NHS Connecting for Health's evaluation of the implementation of the NHS Care Record Service. QRESEARCH undertakes analyses for the Department of Health and other government organisations.

Ethical approval: Trent multicentre research ethics committee.

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

- Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, May M, Brindle P. Derivation and validation of QRISK, a new cardiovascular disease risk score for the United Kingdom: prospective open cohort study. *BMJ* 2007 doi:10.1136/bmj.39261.471806.55.
- Woodward M, Brindle P, Tunstall-Pedoe H. Adding social deprivation and family history to the cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart* 2007;2:172-6.
- National Collaborating Centre for Primary Care. *Section 4.3 of the guideline on cardiovascular risk assessment: the modification of blood lipids for the primary and secondary prevention of cardiovascular disease*. London: NICE, 2008:43.
- Bhopal R, Unwin N, White M, Yallop J, Walker L, Alberti KGMM, et al. Heterogeneity of coronary heart disease risk factors in Indian, Pakistani, Bangladeshi, and European origin populations: cross sectional study. *BMJ* 1999;319:215-20.
- Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. *Heart* 2008;94:34-9.
- Department of Health. *A practical guide to ethnic monitoring in the NHS and social care*. London: DH, 2005:61.
- Anderson KM, Odell PM, Wilson PWF, Kannel WB. Cardiovascular disease risk profiles. *Am Heart J* 1991;121:293-8.
- Maradit-Kremers H, Crowson CS, Nicola PJ, Ballman KV, Roger VL, Jacobsen SJ, et al. Increased unrecognized coronary heart disease and sudden deaths in rheumatoid arthritis: a population-based cohort study. *Arthritis Rheum* 2005;52:402-11.
- Sarnak MJ, Levey AS, Schoolwerth AC, Coresh J, Culleton B, Hamm LL, et al. Kidney disease as a risk factor for development of cardiovascular disease: a statement from the American Heart Association councils on kidney in cardiovascular disease, high blood pressure research, clinical cardiology, and epidemiology and prevention. *Circulation* 2003;108:2154-69.
- Wang TJ, Larson MG, Levy D, Vasan RS, Leip EP, Wolf PA, et al. Temporal relations of atrial fibrillation and congestive heart failure and their joint influence on mortality: the Framingham heart study. *Circulation* 2003;107:2920-5.
- Wolf P, D'Agostino R, Belanger A, Kannel W. Probability of stroke: a risk profile from the Framingham study. *Stroke* 1991;22:312-8.
- National Collaborating Centre for Primary Care. *Cardiovascular risk assessment: the modification of blood lipids for the primary and secondary prevention of cardiovascular disease: full guideline, consultation draft*. London: National Collaborating Centre for Primary Care, 2007:216.
- 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;117:743-53.
- Brindle P, May M, Gill PS, Cappuccio F, D'Agostino Snr R, Fischbacher C, et al. Primary prevention of cardiovascular disease: a web-based risk score for seven British black and minority ethnic groups. *Heart* 2006 doi:10.1136/hrt.2006.092346.
- Nazroo J. South Asian people and heart disease: an assessment of the importance of socioeconomic position. *Ethn Dis* 2001;11:401-11.
- Whitty C, Brunner E, Shipley M, Hemingway H, Marmot M. Differences in biological risk factors for cardiovascular disease between three ethnic groups in the Whitehall II study. *Atherosclerosis* 1999;142:279-86.
- Ranganathan M, Bhopal R. Exclusion and inclusion of nonwhite ethnic minority groups in 72 North American and European cardiovascular cohort studies. *PLoS* 2006;3(3):e44.
- Bhopal R. What is the risk of coronary heart disease in South Asians? A review of UK research. *J Public Health* 2000;22:375-85.

Accepted: 28 May 2008

Fifty years of violent war deaths from Vietnam to Bosnia: analysis of data from world health survey programme

Ziad Obermeyer,^{1,2} Christopher J L Murray,¹ Emmanuela Gakidou¹

EDITORIAL by Garfield

¹Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA

²Harvard Medical School, Boston, MA, USA

Correspondence to: Z Obermeyer
ziad_obermeyer@hms.harvard.edu

BMJ 2008;336:1482-6
doi:10.1136/bmj.a137

ABSTRACT

Objective To provide an accurate estimate of violent war deaths.

Design Analysis of survey data on mortality, adjusted for sampling bias and censoring, from nationally representative surveys designed to measure population health. Estimated deaths compared with estimates in database of passive reports.

Setting 2002-3 World health surveys, in which information was collected from one respondent per household about sibling deaths, including whether such deaths resulted from war injuries.

Main outcome measure Estimated deaths from war injuries in 13 countries over 50 years.

Results From 1955 to 2002, data from the surveys indicated an estimated 5.4 million violent war deaths (95% confidence interval 3.0 to 8.7 million) in 13 countries, ranging from 7000 in the Republic of Congo to 3.8 million in Vietnam. From 1995 to 2002 survey data indicate 36 000 war deaths annually (16 000 to 71 000) in the 13 countries studied. Data from passive surveillance, however, indicated a figure of only a third of this. On the basis of the relation between world health survey data and

passive reports, we estimate 378 000 global war deaths annually from 1985-94, the last years for which complete passive surveillance data were available.

Conclusions The use of data on sibling history from peacetime population surveys can retrospectively estimate mortality from war. War causes more deaths than previously estimated, and there is no evidence to support a recent decline in war deaths.

INTRODUCTION

Estimating mortality due to war is notoriously difficult, but the importance of war as a public health problem as well as a social problem makes it imperative to improve on existing methods of measurement. Accurate estimates of the numbers of deaths are crucial for political, military, and public health planning, as well as for purposes of national history and reconciliation.

Existing techniques

Firstly, a few surveys have been undertaken during wars in specific countries to estimate mortality on the basis of household deaths.¹⁻⁵ These surveys have produced timely estimates, despite the usual problems

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008, doi:10.1136/bmj.a137