

## CARDIOVASCULAR MEDICINE

# Cardiovascular disease risk assessment in older women: can we improve on Framingham? British Women's Heart and Health prospective cohort study

M May, D A Lawlor, P Brindle, R Patel, S Ebrahim



Heart 2006;92:1396-1401. doi: 10.1136/hrt.2005.085381

See end of article for authors' affiliations

Correspondence to: Margaret May, Department of Social Medicine, University of Bristol, Canynge Hall, Whiteladies Road, Bristol BS8 2PR, UK; m.t.may@bristol.ac.uk

Accepted 10 February 2006  
Published Online First 17 March 2006

**Objectives:** To develop a cardiovascular risk assessment tool that is feasible and easy to use in primary care (general practice (GP) model).

**Design:** Prospective cohort study.

**Setting:** 23 towns in the United Kingdom.

**Participants:** 3582 women aged 60 to 79 years who were free of coronary heart disease (CHD) at entry into the British Women's Heart and Health Study.

**Main outcome measures:** Predictive performance of a GP model compared with the standard Framingham model for both CHD and cardiovascular disease (CVD).

**Results:** The Framingham tool predicted CHD events over 5 years accurately (predicted 5.7%, observed 5.5%) but overpredicted CVD events (predicted 10.5%, observed 6.8%). In higher-risk groups, Framingham overpredicted both CHD and CVD events and was poorly calibrated for this cohort. Including C-reactive protein and fibrinogen with standard Framingham risk factors did not improve discrimination of the model. The GP model, which used age, systolic blood pressure, smoking habit and self-rated health (all of which can be easily obtained in one surgery visit) performed as well as the Framingham risk tool: area under the receiver operating curve discrimination statistic was 0.66 (95% confidence interval (CI) 0.62 to 0.70) for CHD and 0.67 (95% CI 0.64 to 0.71) for CVD compared with 0.65 (95% CI 0.61 to 0.68) and 0.66 (95% CI 0.62 to 0.69) for the corresponding Framingham models.

**Conclusions:** An alternative risk assessment based on only a simple routine examination and a small number of pertinent questions may be more useful in the primary care setting. This model appears to perform well but needs to be tested in different populations.

Primary prevention of cardiovascular disease (CVD) involves the identification of people at greatest risk and the targeting of advice and treatments at this group. Most methods use the Framingham score,<sup>1</sup> which requires blood tests and an ECG. Risk scoring has been dominated by the accuracy of model prediction and not by feasibility of application in primary care, where most primary prevention takes place. Furthermore, aetiological factors that may not necessarily be the best prognostic factors or those that are most readily modified have been overemphasised.

Our main objective was to construct a primary care model that would work as well as the Framingham score but would not require an ECG or blood tests. An assessment that requires the primary care professional to undertake only simple routine examinations and ask a few pertinent questions would have obvious advantages over one requiring blood tests and a more time-consuming examination. Such an assessment could be carried out in the surgery with the results, and hence the ability to begin an appropriate prevention plan, immediately available when the patient has been motivated to attend.

A secondary objective was to explore the addition of C-reactive protein and fibrinogen concentrations to the Framingham model, as some authors have claimed that this would improve its performance.<sup>2-4</sup>

## PARTICIPANTS AND METHODS

### The British Women's Heart and Health Study

Full details of the selection of participants and measurements used in the study have been reported.<sup>5,6</sup> Between 1999 and 2001, 4286 women aged 60-79 years who were randomly

selected from 23 British towns were interviewed and examined, completed medical questionnaires and had detailed reviews of their medical records.<sup>5</sup> These women have been followed up over a median of 4.7 years by flagging with the National Health Service (NHS) central register for mortality data and two-yearly review of their medical records. Local ethics committees' approvals were obtained for the study.

### Assessment of prevalent and incident coronary heart disease

Methods used at baseline assessment have been described.<sup>5,6</sup> Prevalent coronary heart disease (CHD) at baseline was defined as a woman with either of the following: (1) a medical record of a myocardial infarction (defined according to World Health Organization criteria), angina, coronary artery bypass or angioplasty; or (2) a self-report that a doctor had ever diagnosed a heart attack or angina. Prevalent stroke and diabetes were similarly defined as a medical record of a diagnosis or a self-report of a doctor's diagnosis. Including both self-reports and data from medical records provides greater assurance that all women with baseline disease have been excluded in the prospective analyses. Incident cases of CHD in women who were free of prevalent disease at baseline were defined as either death with an underlying or contributing cause of CHD (*International classification of diseases*, 10 revision codes I20-I25, I51.6) or a myocardial

**Abbreviations:** AUROC, area under the receiver operating characteristic curve; CHD, coronary heart disease; CVD, cardiovascular disease; GP, general practice

infarction, diagnosis of angina or coronary artery bypass or angioplasty identified in the follow-up medical record review. Incident CVD (CHD or stroke) was similarly defined as freedom from disease at baseline and either death with an underlying or contribution cause coded as I20–I25, I51.6, I60–I69 or G45 or a new CHD or stroke event in any woman's medical record review. The censoring date for events was 31 December 2004.

### Measurement of predictors

Information on smoking and self-rated general health, which was assessed with four prespecified categories—excellent, good, fair and poor—were obtained from the nurse interview or self-completed questionnaire at baseline.<sup>5–6</sup> Blood samples were taken after a minimum 6 h fast. These samples were used to assess lipids by standard procedures.<sup>6</sup> Fibrinogen was assayed in stored citrated plasma by the Clauss assay in an MDA-180 automated coagulometer (Organon Teknika). C-reactive protein was assessed by a high-sensitivity immunonephelometric assay on a ProSpec protein analyser (Dade-Behring) as previously described.<sup>7</sup> Blood pressure, height and weight (used to calculate body mass index) were measured by standard procedures.<sup>6</sup> A resting 12-lead ECG was recorded for each woman at baseline assessment with left ventricular hypertrophy defined according to Minnesota codes (3-1 or 3-3).<sup>8</sup>

### Statistical methods

#### Assessment of the performance of different risk models

We assessed both the calibration and discrimination of the prognostic models.<sup>9</sup> Calibration measures the accuracy of the model predictions and was assessed by comparing mean predicted risk with observed incidence of the outcome overall and in groups of participants classified by level of risk. Discrimination is the ability to rank subjects in order of risk such that those who experience the event of interest have a higher predicted risk than those who do not experience it. If discrimination is poor, then people above the threshold will receive treatment unnecessarily and people below the threshold will experience CVD but will not receive preventive treatment. If calibration is poor then the risk thresholds for treatment will be set higher or lower than they should be to achieve the primary prevention targets.

#### Test of the predictions from the published Framingham models

The published Framingham equations<sup>1</sup> for CHD and CVD events were used to predict risk of an event occurring during the time between the baseline survey and the first follow up for each woman. The predicted risk was used to rank the women and classify them by fifths of risk. The mean predicted risk for each fifth and the overall risk were compared with the corresponding observed number of events. The ratio of predicted risk to observed incidence was calculated to see whether the Framingham score overpredicted or underpredicted risk in our study. Receiver operating characteristic analysis, plotting sensitivity and 1 – specificity of risk predictions (see appendix), was used to assess the discrimination of the risk score.

#### Comparison between the general practice model and the Framingham model

Weibull proportional hazards survival models and standard methods for assessing calibration and accuracy of model predictions were used to see whether non-Framingham risk factors for CVD could be used either in addition to or instead of the standard Framingham risk factors. Multiple imputation methods were used to deal with missing data,

**Table 1** Risk factors for coronary heart disease in 3582 women

Risk factor	Mean (SD)	No (%) imputed
Age (years)	68.6 (5.5)	1 (0.03%)
Systolic blood pressure (mm Hg)	148 (25)	260 (7%)
Total cholesterol	6.6 (1.2)	360 (10%)
HDL cholesterol	1.7 (0.5)	360 (10%)
Log total:HDL cholesterol	1.4 (0.3)	360 (10%)
Body mass index (kg/m <sup>2</sup> )	27.4 (4.9)	263 (7%)
Waist measurement (cm)	85.8 (12.0)	275 (8%)
Log C-reactive protein	0.55 (1.14)	366 (10%)
Fibrinogen	3.4 (0.7)	322 (9%)
	<b>No (%)</b>	
Left ventricular hypertrophy	379 (11%)	0
Diabetes	157 (4%)	0
Smoking		15 (0.4%)
Never smoker	2010 (56%)	
Former smoker	1137 (32%)	
Current smoker (cigarettes/day)		
0–9	163 (5%)	
≥10	258 (7%)	
Self-rated health		186 (5%)
Excellent	405 (12%)	
Good	2172 (64%)	
Fair	749 (22%)	
Poor	71 (2%)	

HDL, high density lipoprotein.

facilitating comparisons across the different models (see appendix).

### RESULTS

After removal of women with baseline CHD or CVD, 3582 were available for the analyses with CHD (704 women with baseline disease removed) as the outcome and 3511 with CVD (775 women with baseline disease removed) as the outcome. Table 1 shows the mean (SD) (or proportions) of the risk factors in the 3582 women together with the number of women with imputed data.

No variables had more than 11% imputed data and the five imputed data sets were statistically similar to the actual data. Results for the analyses including imputed data were similar to the complete case analyses.

#### Test of the predictions from the published Framingham models

Table 2 shows the predicted Framingham risk and observed incidence of CHD and CVD events with the ratio of predicted to observed for fifths of risk and overall.

Follow-up time was an average of 4.7 years (range 3.4–5.7 years). For CHD the predicted risk of an event was 5.7% and the observed incidence was 5.5% (198 cases), an overprediction of 3%. However, the Framingham score underpredicted in the low-risk fifths and overpredicted in the highest-risk fifth. For CVD, the predicted risk of an event was 10.5% and the observed incidence was 6.8% (240 cases), an overprediction of 54%. The overprediction was greatest in the two highest-risk fifths. The discrimination (area under the receiver operating characteristic curve (AUROC)) of the Framingham model was 0.59 for CHD and 0.62 for CVD events when the women were classified by fifths of risk and 0.63 and 0.64, respectively, for the ranked risk.

#### Comparison of the discrimination of candidate risk prediction models

Table 3 shows the hazard ratios (95% confidence intervals (CIs)) for CHD and CVD events comparing a model with the standard Framingham risk factors with one that also included log C-reactive protein and fibrinogen. In this dataset, C-reactive protein and fibrinogen predicted CHD in

**Table 2** Predicted Framingham risk and observed incidence of CHD and CVD events by fifths of risk

Fifth of risk	No of subjects	No of events	Predicted (%)	Observed (%)	Ratio predicted to observed
<b>CHD event</b>					
1	717	21	1.56	2.93	0.53
2	716	24	2.94	3.35	0.88
3	717	31	4.34	4.32	1.00
4	716	57	6.43	7.96	0.81
5	716	65	13.22	9.08	1.46
Overall	3582	198	5.69	5.53	1.03
AUROC					
Using fifths			0.59 (95% CI 0.56 to 0.63)		
Using ranked data (maximum)			0.63 (95% CI 0.59 to 0.67)		
Sensitivity/specificity					
30% 10-year CHD risk threshold				10%/95%	
15% 10-year CHD risk threshold				44%/74%	
<b>CVD event</b>					
1	703	22	3.14	3.13	1.00
2	702	32	6.09	4.56	1.34
3	702	42	8.88	5.98	1.48
4	702	55	12.58	7.83	1.61
5	702	89	21.83	12.68	1.72
Overall	3511	240	10.50	6.84	1.54
AUROC					
Using fifths			0.62 (95% CI 0.58 to 0.65)		
Using ranked data (maximum)			0.64 (95% CI 0.61 to 0.68)		
Sensitivity/specificity					
30% 10-year CVD risk threshold				38%/79%	
15% 10-year CVD risk threshold				85%/30%	

AUROC, area under the receiver operating characteristic curve; CHD, coronary heart disease; CVD, cardiovascular disease.

univariable models, but the hazard ratios were attenuated in the multivariable model. Including C-reactive protein and fibrinogen did not improve discrimination for either outcome.

The proposed general practice (GP) model included the standard risk factors of age, systolic blood pressure and

smoking status but not cholesterol ratio, diabetes and left ventricular hypertrophy, as these require laboratory tests or an ECG. The alternative risk factors considered for inclusion in the GP model were body mass index or waist measurement and self-rated health (excellent, good, fair or poor). In contrast with the published Framingham model, a four-category

**Table 3** Hazard ratios (95% CI) for CHD and CVD events estimated from models fitted on BWHHS data using Framingham risk factors (model 1) and with addition of C-reactive protein and fibrinogen (model 2)

	Model 1	Model 2
<b>CHD event</b> (n = 3582 women; n = 198 events)		
Age (per 5 years)	1.24 (1.08 to 1.42)	1.24 (1.08 to 1.42)
Systolic blood pressure (per 10 mm Hg)	1.06 (0.99 to 1.13)	1.05 (0.99 to 1.13)
Total:HDL cholesterol ratio (log)	2.26 (1.42 to 3.60)	2.10 (1.31 to 3.37)
Current smoker	2.35 (1.68 to 3.30)	2.26 (1.60 to 3.19)
Left ventricular hypertrophy	1.27 (0.84 to 1.93)	1.24 (0.82 to 1.89)
Diabetes	0.90 (0.46 to 1.77)	0.87 (0.44 to 1.71)
C-reactive protein (log)		1.09 (0.94 to 1.27)
Fibrinogen		1.04 (0.82 to 1.31)
AUROC using rank	0.64 (0.60 to 0.69)	0.64 (0.60 to 0.69)
P value for difference in AUROCs = 0.76		
Sensitivity/specificity		
30% CHD risk threshold	5%/99%	6%/99%
15% CHD risk threshold	30%/85%	31%/86%
<b>CVD event</b> (n = 3511 women; n = 240 events)		
Age (per 5 years)	1.28 (1.13 to 1.45)	1.28 (1.13 to 1.45)
Systolic blood pressure (per 10 mm Hg)	1.04 (0.98 to 1.11)	1.04 (0.98 to 1.10)
Total:HDL cholesterol ratio (log)	2.20 (1.42 to 3.43)	2.04 (1.29 to 3.21)
Current smoker	2.65 (1.96 to 3.59)	2.55 (1.87 to 3.46)
Left ventricular hypertrophy	1.35 (0.92 to 1.97)	1.32 (0.90 to 1.93)
Diabetes	1.06 (0.59 to 1.90)	1.02 (0.56 to 1.83)
C-reactive protein (log)		1.11 (0.96 to 1.27)
Fibrinogen		1.03 (0.83 to 1.28)
AUROC using rank	0.65 (0.62 to 0.69)	0.66 (0.62 to 0.69)
P value for difference in AUROCs = 0.63		
Sensitivity/specificity		
30% CVD risk threshold	12%/97%	13%/97%
15% CVD risk threshold	46%/75%	46%/74%

AUROC, area under the receiver operating characteristic curve; BWHHS, British Women's Heart and Health Study; CHD, coronary heart disease; CVD, cardiovascular disease; HDL, high density lipoprotein.

**Table 4** Hazard ratios (95% CI) for CHD and CVD events for each of four GP models

	Model 1	Model 2	Model 3	Model 4
<b>CHD event</b> (n = 3582 women; n = 198 events)				
Age (per 5 years)	1.24 (1.08 to 1.42)	1.19 (1.04 to 1.36)	1.25 (1.09 to 1.43)	1.20 (1.05 to 1.37)
SBP (per 10 mm Hg)	1.07 (1.00 to 1.14)	1.06 (1.00 to 1.13)	1.06 (1.00 to 1.13)	1.06 (0.99 to 1.13)
Smoking (v never smoker)				
Former smoker	1.38 (1.00 to 1.90)	1.32 (0.96 to 1.82)	1.35 (0.98 to 1.86)	1.30 (0.94 to 1.80)
Current 0–9 cigs/day	2.43 (1.42 to 4.14)	2.26 (1.31 to 3.88)	2.46 (1.44 to 4.20)	2.32 (1.35 to 3.99)
Current ≥10 cigs/day	3.06 (2.00 to 4.68)	2.56 (1.66 to 3.95)	3.01 (1.97 to 4.60)	2.60 (1.69 to 4.00)
Body mass index (per kg/m <sup>2</sup> )	1.01 (0.98 to 1.04)	0.99 (0.97 to 1.02)		
Waist (per 10 cm)			1.14 (1.02 to 1.28)	1.08 (0.96 to 1.21)
Self-rated health (v excellent)				
Good		1.60 (0.82 to 3.11)		1.55 (0.79 to 3.03)
Fair		2.69 (1.36 to 5.31)		2.51 (1.27 to 4.99)
Poor		7.40 (3.22 to 17.01)		6.79 (2.95 to 15.64)
GP model AUROC	0.63 (0.59 to 0.67)	0.66 (0.62 to 0.70)	0.63 (0.59 to 0.67)	0.66 (0.62 to 0.70)
P value for difference from Framingham model AUROC	0.09	0.48	0.16	0.56
Sensitivity/specificity				
30% CHD risk threshold	4%/99%	11%/98%	5%/99%	6%/99%
15% CHD risk threshold	26%/86%	34%/86%	27%/86%	34%/85%
<b>CVD event</b> (n = 3511 women; n = 240 events)				
Age (per 5 years)	1.29 (1.14 to 1.46)	1.24 (1.09 to 1.40)	1.30 (1.14 to 1.47)	1.25 (1.10 to 1.41)
SBP (per 10 mm Hg)	1.06 (1.00 to 1.12)	1.05 (0.99 to 1.11)	1.05 (0.99 to 1.11)	1.05 (0.99 to 1.11)
Smoking (v never smoker)				
Former smoker	1.27 (0.95 to 1.71)	1.23 (0.91 to 1.65)	1.25 (0.93 to 1.67)	1.21 (0.90 to 1.62)
Current 0–9 cigs/day	2.63 (1.65 to 4.21)	2.47 (1.53 to 3.97)	2.65 (1.65 to 4.23)	2.52 (1.56 to 4.05)
Current ≥10 cigs/day	3.33 (2.28 to 4.87)	2.86 (1.95 to 4.21)	3.28 (2.25 to 4.79)	2.90 (1.97 to 4.25)
Body mass index (per kg/m <sup>2</sup> )	1.01 (0.99 to 1.04)	1.00 (0.97 to 1.02)		
Waist (per 10 cm)			1.15 (1.04 to 1.28)	1.09 (0.98 to 1.21)
Self-rated health (v excellent)				
Good		1.81 (0.96 to 3.41)		1.76 (0.94 to 3.32)
Fair		2.97 (1.56 to 5.66)		2.78 (1.45 to 5.33)
Poor		9.33 (4.27 to 20.41)		8.66 (3.93 to 19.09)
GP model AUROC	0.64 (0.60 to 0.67)	0.67 (0.64 to 0.71)	0.64 (0.60 to 0.68)	0.67 (0.63 to 0.71)
P value for difference from Framingham model AUROC	0.10	0.31	0.23	0.32
Sensitivity/specificity				
30% CVD risk threshold	10%/97%	18%/96%	11%/97%	11%/97%
15% CVD risk threshold	43%/74%	48%/76%	47%/73%	48%/74%

Area under the receiver operating characteristic curve (AUROC) (95% CI) for Framingham models fitted to BWHHS (British Women's Heart and Health Study) data: coronary heart disease (CHD), 0.65 (0.61 to 0.69); cardiovascular disease (CVD), 0.66 (0.62 to 0.69). cigs, cigarettes; GP, general practice; SBP, systolic blood pressure.

classification of smoking was used in the GP model: never smoker, former smoker and current smoker of either 0–9 or 10 or more cigarettes a day. Self-rated health was a particularly strong predictor of events with a hazard ratio for “poor” compared with “excellent” of 9.6 (95% CI 4.1 to 22.9) for CHD and 11.4 (95% CI 5.1 to 25.6) for CVD. Body mass index was not an independent predictor of CHD or CVD.

Table 4 shows the hazard ratios (95% CI) for the prognostic factors in the GP models for both outcomes with the AUROCs and the p value for the difference in discrimination of the models compared with the Framingham model. Discrimination appeared to be marginally better with GP model 2, but CIs of the AUROCs for all GP models and for the corresponding Framingham model overlap.

Tables 2–4 also show the sensitivity and specificity at 30% and 15% 10-year risk thresholds for the published Framingham model and for the models estimated on this cohort, adjusted for shorter follow-up time. As the sensitivity and specificity of the published Framingham equations are different from those estimated in the model with Framingham risk factors fitted to this cohort, the Framingham risk score is not well calibrated to this population at the important treatment decision-making thresholds.

## DISCUSSION

In a group of older British women we found that a simple risk assessment based on age, systolic blood pressure, smoking habit and self-rated health performed as well as the

Framingham risk assessment. The addition of C-reactive protein or fibrinogen did not improve the performance of the Framingham equation.<sup>10–12</sup> The Framingham risk assessment overpredicted, particularly for CVD, and particularly in those at higher risk. This pattern has been shown in a number of other studies of different populations in which the Framingham equation has been assessed<sup>13 14</sup> and may result from variations in the performance of the equation in different populations or overfit of the model to the data in which it was originally developed.<sup>15</sup> Owing to regression to the mean, one would anticipate the pattern of overprediction in high-risk and underprediction in low-risk groups when a prediction model generated in one dataset is applied to an independent dataset.

## Limitations of this study

Variables for inclusion in the GP model were selected a priori. As the model is being used to predict in the same dataset that was also used to estimate the coefficients, predictive performance is likely to be overoptimistic—that is, the discrimination or the calibration of the model may be expected to worsen if it were used for prediction in independent data.<sup>16</sup> Further work is required to produce a model that would generalise well to all British women, as well as other populations, including validation of any such model in the primary care setting.<sup>17</sup> Prediction models, like any screening test, should be evaluated in randomised trials to determine whether they are effective.

Ideally, assessment of 10-year risk of CVD events should be based on 10 years of follow up, but only five years of follow up was available in our study. This probably did not bias the results substantially, as follow-up time was taken into account in the statistical methods. However, events might have been somewhat underascertained close to the end of follow up, which might have artificially contributed to the overprediction of events by the Framingham equations. However, for CHD it was the distribution of events over the fifths of risk that did not match, rather than the total number of events. CVD events were particularly overpredicted in our study. Results could have been affected by differences in event definition and ascertainment in Framingham compared with this cohort. Furthermore, the Framingham risk score was not designed for patients more than 74 years old, although primary prevention in this age group, and therefore our attempt to develop an easy to use and accurate prediction tool for them, is important.

Risk factors measured in epidemiological studies may not adequately reflect the methods of measurement in routine primary care, where there may be more repeated measurements but less standardised measurement. Our model depends on the validity of the self-rated health question. Research from the British Regional Heart Study has indicated that it is a reliable instrument.<sup>18</sup> Self-rated health in the context of opportunistic screening when someone presents to their GP with symptoms in particular may be very different from our assessment in this study. However, in primary prevention screening clinics, where patients are not presenting because of ill health, we expect our model would perform in a similar way to our study.

### Implications

Most work to date on the performance of prediction tools has concentrated on accuracy rather than feasibility. The focus of the modelling has not been on prognostic factors, but rather on aetiological factors. Our findings suggest that a simple model based on age, systolic blood pressure, smoking habit and self-rated health performs well for the prediction of CVD in older British women. Further research is required to examine the performance of this model in other independent datasets and different populations, but if it is found to perform well in these studies it may improve risk detection and primary prevention. The Framingham tool has not been widely used,<sup>19</sup> possibly because it requires blood tests and ECG assessment. A tool that requires only simple tests and a few questions may be more widely used. Our GP model has the advantage over a model requiring blood tests and ECG that it would provide immediate results. Thus, primary prevention advice and treatment can begin sooner and do not require the patient to return for a second visit to discuss their results (with the risk that some patients will not return for this). On the other hand, risk assessment based on blood tests and ECG may be a more powerful motivator to comply with treatment. Ultimately, a randomised trial is required to compare the costs and effects of the different methods on primary prevention.

The Framingham equations are not well calibrated to this population of elderly British women, particularly in high-risk groups, where the decision thresholds for treatment lie, nor do they have high discrimination. If the Framingham tool continues to be used, then the threshold for treatment specified in the Department of Health guidelines should be revised downwards for older women to increase sensitivity of the risk assessment.

The nature of the prognostic variables used in prediction models will, to a certain extent, drive the focus of the intervention offered to patients found to be at high risk. For example, if cholesterol is measured and found to be high,

then statins may be prescribed, but if a model uses body mass index as a predictor of CHD, then advice on lifestyle changes in diet and exercise are more likely to be given. Research on the acceptability, effectiveness and cost of different interventions in this population can inform the selection of prognostic variables for a more appropriate risk assessment tool. Furthermore, on economic grounds, using a two-stage screening process would be advantageous. A cheap screening tool such as the GP model can be used to screen out those patients with low risk. In the second stage, the more expensive and invasive blood tests can be offered to the remaining patients to determine who is at highest risk and would benefit most from drug treatments.

### ACKNOWLEDGEMENTS

We thank all of the general practitioners and their staff who supported data collection and the women who participated in the study.

### Authors' affiliations

**M May, D A Lawlor, P Brindle, R Patel**, Department of Social Medicine, University of Bristol, Bristol, UK

**S Ebrahim**, Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK

The British Women's Heart and Health Study is funded by the UK Department of Health and British Heart Foundation. MM is funded by the British Heart Foundation and the Medical Research Council. DAL is funded by a UK Department of Health Career Scientist Award. The funding bodies have not influenced any aspects of the study design, analysis or interpretation of results. The views expressed in this publication are those of the authors and not necessarily those of any of the funding bodies.

Competing interests: None declared.

Contributions: Margaret May developed the study aim, undertook all statistical analyses and wrote the first draft of the paper. Debbie Lawlor co-directs the British Women's Heart and Health Study, thought of and developed the study aim and contributed to writing the paper. Peter Brindle developed the study aim and contributed to writing the paper. Rita Patel managed the British Women's Heart and Health Study database and contributed to writing the paper. Shah Ebrahim is the principal investigator of the British Women's Heart and Health Study, thought of and developed the study aim and contributed to writing the paper.

### REFERENCES

- Anderson KM, Odell PM, Wilson PWF. Cardiovascular disease risk profiles. *Am Heart J* 1990;**121**:293–8.
- Danesh J, Collins R, Appleby P, et al. Association of fibrinogen, C-reactive protein, albumin, or leukocyte count with coronary heart disease: meta-analyses of prospective studies. *JAMA* 1998;**279**:1477–82.
- Koenig W, Lowel H, Baumert J, et al. C-reactive protein modulates risk prediction based on the Framingham Score: implications for future risk assessment: results from a large cohort study in southern Germany. *Circulation* 2004;**109**:1349–53.
- Kuller LH, Tracy RP, Shaten J, et al. Relation of C-reactive protein and coronary heart disease in the MRFIT nested case control study. Multiple Risk Factor Intervention Trial. *Am J Epidemiol* 1996;**144**:537–47.
- Lawlor DA, Bedford C, Taylor M, et al. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J Epidemiol Community Health* 2003;**57**:134–40.
- Lawlor DA, Davey Smith G, Ebrahim S. Life course influences on insulin resistance: findings from the British Women's Heart and Health Study. *Diabetes Care* 2003;**26**:97–103.
- Wannamethee SG, Lowe GD, Whincup PH, et al. Physical activity and hemostatic and inflammatory variables in elderly men. *Circulation* 2002;**105**:1785–90.
- Prineas RJ, Crow RS, Blackburn H. *The Minnesota manual of electrographic findings*. Bristol: John Wright Publications, 1982.
- Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;**130**:515–24.
- Lawlor DA, Davey Smith G, Rumley A, et al. Associations of fibrinogen and C-reactive protein with prevalent and incident coronary heart disease are attenuated by adjustment for confounding factors: British Women's Heart and Health study. *Thromb Haemost* 2005;**93**:955–63.
- Davey Smith G, Lawlor DA, Harbord R, et al. Association of C-reactive protein with blood pressure and hypertension: life course confounding and

- Mendelian randomization tests of causality. *Arterioscler Thromb Vasc Biol* 2005;**25**:1051–6.
- 12 **Davey Smith G**, Harbord R, Ebrahim S. Fibrinogen, C-reactive protein and coronary heart disease: does Mendelian randomization suggest the associations are non-causal? *QJM* 2004;**97**:163–6.
  - 13 **Empana JP**, Ducimetiere P, Arveiler D, *et al*. Are the Framingham and PROCAM coronary heart disease risk functions applicable to different European populations? The PRIME study. *Eur Heart J* 2003;**24**:1903–11.
  - 14 **Diverse Populations Collaborative Group**. Prediction of mortality from coronary heart disease among diverse populations: is there a common predictive function? *Heart* 2002;**88**:222–8.
  - 15 **Brindle P**, Emberson J, Lampe F, *et al*. Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. *BMJ* 2003;**327**:1267–70.
  - 16 **Efron B**. How biased is the apparent error rate of a prediction rule? *J Am Stat Assoc* 1986;**81**:461–70.
  - 17 **May M**, Royston P, Egger M, *et al*. Development and validation of a prognostic model for survival time data: application to prognosis of HIV positive patients treated with antiretroviral therapy. *Stat Med* 2004;**23**:2375–98.
  - 18 **Wannamethee G**, Shaper AG. Self-assessment of health status and mortality in middle-aged British men. *Int J Epidemiol* 1991;**20**:239–45.
  - 19 **Montgomery AA**, Fahey T, Peters TJ, *et al*. Evaluation of computer based clinical decision support system and risk chart for management of hypertension in primary care: randomised controlled trial. *BMJ* 2000;**320**:686–90.
  - 20 **Van Buuren S**, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Stat Med* 1999;**18**:681–94.
  - 21 **Royston P**. Multiple imputation of missing values. *Stata J* 2004;**4**:227–41.
  - 22 **Rubin DB**. *Multiple imputation for nonresponse in surveys*. New York: Wiley, 1987.

## APPENDIX: STATISTICAL METHODS

### 1. MULTIPLE IMPUTATION OF MISSING DATA

To reduce bias and to allow comparisons of regression models based on all subjects, we used multiple imputation of missing data with switching regression<sup>20, 21</sup> rather than complete case analysis. All prognostic variables, log of the survival times for coronary heart disease (CHD) and cardiovascular disease (CVD), and both censoring indicators were used in the

imputation regressions. After imputation, five imputed data sets were analysed and the results combined appropriately by using Rubin's rules.<sup>22</sup> Women with prevalent CHD at baseline were excluded from all regression analyses and those with a previous diagnosis of stroke were also excluded from analyses with CVD as the end point. The point estimates were similar when the analysis was repeated on the subset of subjects with complete data.

### 2. ASSESSMENT OF DISCRIMINATION BY RECEIVER OPERATING CHARACTERISTIC ANALYSIS

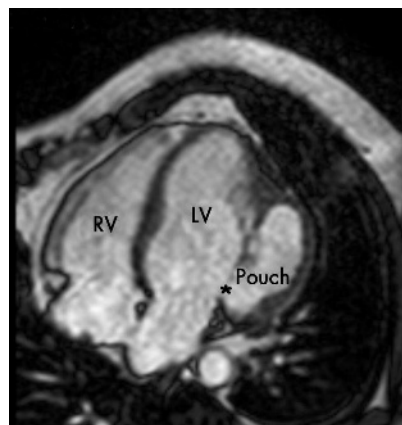
The receiver operating characteristic (ROC) curve is a graph of sensitivity versus 1 – specificity of the classification. The area under the ROC curve (AUROC) can be interpreted as the probability that a randomly selected subject who experiences the event has a higher predicted risk than a randomly selected person who does not experience the event. Overall discrimination of the Framingham models was assessed with the AUROC. As an area of 0.5 would be expected by chance allocation, a good prediction model would have an AUROC of 0.75 or higher. For comparison, the maximum possible AUROC for these data and models, calculated with the ranked risk scores and thus independent of any choice of classification threshold, was determined.

We calculated the AUROC on the ranked risk scores predicted by the candidate models and compared them with the AUROC of the model that included only the standard Framingham risk factors. To determine how well the published Framingham equations and the candidate Weibull prognostic models could classify the women into high and low-risk groups, we calculated the sensitivity and specificity of the model predictions for both CHD and CVD events at 30% and 15% 10-year risk thresholds.

## IMAGES IN CARDIOLOGY .....

doi: 10.1136/hrt.2005.079699

### Double chambered left ventricle in cardiac magnetic resonance imaging



**A** 27-year-old man was referred for evaluation of right-sided thoracic pain. Physical examination revealed no cardiac murmur. ECG showed left bundle branch block. At transthoracic echocardiography an additional chamber, attached to the left ventricle, was noted.

For further investigation cardiac magnetic resonance imaging with black blood sequences and multiphased SSFP sequences were performed. In all projections we saw a ventricle-like "pouch", appending to the posterior and lateral wall of the left ventricle (panel shows four chamber view, with a septal defect\*). Its size was 5.1 cm in long axis and 2.0 cm in short axis. The pouch had no inflow portion and was not part of the apex. The wall thickness of the pouch was 8 mm in end-diastole and equal to the remaining left ventricular myocardium. The septum separating the pouch from the regular left ventricle was interrupted by two septal defects of 1.9 cm and 0.5 cm in size.

In the cine sequences we saw a slightly reduced left ventricular contraction (ejection fraction 46%) (to view video files 1 and 2 visit the *Heart* website—<http://www.heartjnl.com/> supplemental). There was normal systolic contraction of the pouch without any regional wall motion abnormality.

After injection of Gd-DTPA the first pass perfusion in a short axis view showed normal myocardial perfusion. No evidence for myocardial scarring in the late enhancement in short and long axis could be demonstrated.



To view video footage visit the *Heart* website—<http://www.heartjnl.com/supplemental>

T Hemmers, M Schwaiger, H Stern  
hemmers@dhn.mhn.de