The Framingham and UK Prospective Diabetes Study (UKPDS) risk equations do not reliably estimate the probability of cardiovascular events in a large ethnically diverse sample of patients with diabetes: the Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation (ADVANCE) Study

A. P. Kengne · A. Patel · S. Colagiuri · S. Heller ·
P. Hamet · M. Marre · C. Y. Pan · S. Zoungas ·
D. E. Grobbee · B. Neal · J. Chalmers · M. Woodward ·
for the ADVANCE Collaborative Group

Abstract

Aims/hypothesis Available multivariable equations for cardiovascular risk assessment in people with diabetes have been derived either from the general population or from populations with diabetes. Their utility and comparative performance in a contemporary group of patients with type 2 diabetes are not well established. The aim of this study was to evaluate the performance of the Framingham and UK Prospective Diabetes Study (UKPDS) risk equations in participants who took part in the Action in Diabetes and Vascular disease: Preterax and Diamicron-MR Controlled Evaluation (ADVANCE) trial.

Methods The 4-year risks of cardiovascular disease (CVD) and its constituents were estimated using two published

A. P. Kengne (✉) · A. Patel · S. Zoungas · B. Neal ·
J. Chalmers · M. Woodward
The George Institute for International Health,
University of Sydney,
PO Box M201, Level 10 KGV Building, RPAH, Missenden Road,
Camperdown,
Sydney, NSW 2050, Australia
e-mail: akengne@george.org.au

S. Colagiuri
Institute of Obesity, Nutrition and Exercise, Faculty of Medicine,
The University of Sydney,
Sydney, NSW, Australia

S. Heller
University of Sheffield and Sheffield Teaching Hospitals National
Health Service (NHS) Foundation Trust,
Sheffield, UK

P. Hamet
Centre Hospitalier de l'Université de Montréal,
Université de Montréal,
Montréal, QC, Canada

M. Marre
Hôpital Bichat-Claude Bernard and Université Paris 7,
Paris, France

C. Y. Pan
Chinese People's Liberation Army General Hospital,
Beijing, China

S. Zoungas
School of Public Health, Monash University,
Melbourne, Victoria, Australia

D. E. Grobbee
Julius Centre for Health Sciences and Primary Care,
University Medical Centre Utrecht,
Utrecht, the Netherlands

M. Woodward
Mount Sinai School of Medicine,
New York, NY, USA

Framingham and the UKPDS risk equations in 7,502 individuals with type 2 diabetes without prior known CVD at their enrolment in the trial.

Results The risk of major CVD was overestimated by 170% (95% CI 146–195%) and 202% (176–231%) using the two Framingham equations. The risk of major coronary heart disease was overestimated by 198% (162–238%) with the UKPDS, and by 146% (117–179%) and 289% (243–341%) with the two different Framingham equations, respectively. The risks of stroke events were also overestimated with the UKPDS and one of the Framingham equations. The ability of these equations to rank risk among ADVANCE participants was modest, with *c*-statistics ranging from 0.57 to 0.71. Results stratified by sex, treatment allocation and ethnicity were broadly similar.

Conclusions/interpretation Application of the Framingham and UKPDS risk equations to a contemporary treated group of patients with established type 2 diabetes is likely to substantially overestimate cardiovascular risk.

Keywords Calibration · Cardiovascular disease · Diabetes mellitus · Discrimination · Risk prediction

Abbreviations

ADVANCE	Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation
CVD	Cardiovascular disease
ROC	Receiver-operating characteristic curve
UKPDS	UK Prospective Diabetes Study

Introduction

The determinants of cardiovascular disease (CVD) are multifactorial. A major advance in the field of CVD prevention over the past four decades has been the development of statistical models to predict the probability of future cardiovascular events on the basis of multiple risk factors [1, 2]. While some suggest that the presence of diabetes should be regarded as a risk equivalent to established CHD, obviating the need for risk stratification models, others have developed and advocated the use of such tools in people with diabetes [3–5]. Predictors included in existing models vary in their number as well as in how they are treated in the models [6]. Some risk prediction equations, including those derived from the Framingham Study, have treated diabetes status as just one of many prognostic variables [7, 8]. The reliability of the Framingham equations for predicting risk of CVD in people with diabetes has been questioned [9], as the original equation was derived from a population with only

337 diabetic patients [10] and included no specific measure of glycaemia.

Risk prediction equations have also been developed specifically for individuals with diabetes [9, 11–13]. The best known are those of the UK Prospective Diabetes Study (UKPDS) [9, 11]. A particular feature of some of these models is that they consider glucose measurements and the duration of diagnosed diabetes. Type 2 diabetes is characterised by its insidious course and potentially long asymptomatic period. Given variations in health services and clinical practice, the extent to which time since diagnosis represents cumulative exposure to higher than optimal glucose levels may vary substantially across populations. This may affect the applicability of the UKPDS models and other diabetes-specific risk models to different populations. In addition, secular changes in the management of diabetes since the UKPDS Study was completed may affect the current applicability of models derived from that cohort.

In these analyses, we assess the performance of the UKPDS model and commonly used general Framingham CVD risk models in a broad, contemporary group of patients with type 2 diabetes, who participated in the Action in Diabetes and Vascular Disease: Preterax and Diamicron-MR Controlled Evaluation (ADVANCE) Study. The trial was registered with ClinicalTrial.gov (trial no. NCT00145925).

Methods

The protocol for ADVANCE has been published in detail [14, 15]. In brief, ADVANCE was a 2×2 factorial randomised controlled trial of blood pressure lowering (perindopril–indapamide vs placebo) and glucose control (glicazide MR-based intensive intervention vs standard care) among 11,140 individuals with type 2 diabetes recruited from 215 centres across 20 countries in Australasia, Asia, Europe and North America. Participants had to be at least 55 years of age at entry, to have been diagnosed with diabetes at the age of 30 years or older and to be at further increased risk of vascular events. For the current analyses, only those participants were included who were free of known CVD at baseline and for whom complete data for computation of global cardiovascular risk according to the UKPDS and Framingham risk equations were available.

Baseline assessment At recruitment, baseline data were collected on medical history, current medical treatment and several risk factors, including blood glucose, HbA_{1c}, lipids, blood pressure, BMI, atrial fibrillation, smoking and history of CVD.

Outcomes The following primary outcomes were considered: major CHD events, major cerebrovascular events

and major CVD events. Major CHD events included death from CHD, sudden death and non-fatal myocardial infarction. Major cerebrovascular events included death from cerebrovascular disease and non-fatal stroke. Major CVD included any CVD death, non-fatal myocardial infarction and non-fatal stroke. Three secondary outcomes were also defined: (1) any CHD event (including any major CHD, coronary revascularisation and hospitalisation for unstable angina); (2) any cerebrovascular event (including any stroke and transient ischaemic attack); and (3) total CVD events (including any of the above, congestive heart failure and peripheral vascular disease). Outcomes were coded according to the 10th revision of the International Classification of Diseases (ICD-10; www.who.int/classifications/icd/en/) and major events (suspected myocardial infarction, suspected stroke and all deaths) were centrally validated by an independent endpoint committee. Analyses were restricted to the first relevant cardiovascular event during follow-up.

Cardiovascular risk models The baseline characteristics of the ADVANCE trial participants were used to calculate the expected 4-year probability of CHD and cerebrovascular events using the UKPDS risk equations [9, 11] and two different Framingham equations, namely Anderson et al. [8] and D'Agostino et al. [16]. In addition, the expected probability of total CVD was also derived from Framingham equations [8, 16]. There is no publicly available UKPDS risk equation for total CVD events. The Framingham Anderson equations have been recommended for risk estimation over a range of 4 to 12 years [8]. The UKPDS equations have been designed for risk evaluation for any duration of follow-up (t) in years in a patient who has been newly diagnosed with type 2 diabetes or who has had diabetes for a known duration of years (T) [9, 11]. The Framingham D'Agostino equations were published with the 10-year baseline survival rates [16]. For these analyses, the 4-year baseline survival for men (0.95964) and women (0.98215) was obtained from the Framingham investigators. This approach assumes that the regression coefficients are constant over time.

Assessment of model performance Discrimination, i.e. the model's ability to rank individual risk, was evaluated using the area under the receiver-operating characteristic curve (ROC, also referred to as AUC) and non-parametric methods [17]. Calibration, i.e. the agreement between expected (E) and observed (O) rates, was assessed overall and within pre-specified groups of participants. The 95% CIs for the expected:observed ratio were estimated by assuming Poisson variance [18]. The overall bias in expected event rates was estimated as $100(E-O)/O$ [19]. The difference in expected and observed probabilities of an event within pre-specified groups of participants (ranked according to fifths of predicted probability) was tested

using the Hosmer–Lemeshow χ^2 statistic (HLχ^2) [20]. We also assessed whether recalibration, i.e. adjusting the models by replacing average values of predictors and event rates in the original population by those in the target population (i.e. the ADVANCE Study), could improve performance of the risk models [21–24]. Recalibration was restricted to the outcomes major CVD, major CHD and stroke. The recalibration method proposed by the Framingham investigators was used to recalibrate the Framingham D'Agostino equations [21, 22]; and the method proposed by van Houwelingen [24] was used for the Framingham Anderson and UKPDS equations, which are both parametric models. The UKPDS equations were originally modelled at the mean levels of continuous predictors (age, systolic BP, HbA$_{1c}$ and total cholesterol:HDL-cholesterol ratio) specific to the UKPDS population [9, 11]. Therefore, the UKPDS equations were also recalibrated by replacing the average levels of predictors and intercept terms in the original UKPDS equations [9, 11] with those observed for the ADVANCE population. The intercept q_0 in the UKPDS equations has been described as the approximate probability of CHD ($q_0 = 0.0112$) or stroke ($q_0 = 0.00186$) in the 1st year of diagnosed diabetes [9, 11]. The life-table method was used, with a single year as time unit, to derive the probability of major CHD ($q_0 = 0.00757$) and stroke ($q_0 = 0.00190$) events during the 1st year into the study for the 530 participants (52% men) who were diagnosed with diabetes less than 1 year before their enrolment in the ADVANCE study.

Main analyses were conducted for those members of the overall ADVANCE cohort who were free of CVD at baseline. To explore any effect of randomised treatment, two sensitivity analyses were conducted: (1) using only the 'double control' group, i.e. those assigned to the placebo group in the blood pressure lowering arm and the standard care group of the blood glucose control arm ($n = 1,882$); and (2) by replacing the baseline values of blood pressure and HbA$_{1c}$ with the values at 1 year of follow-up for participants who did not experience a cardiovascular event during this period. Analyses were also conducted after stratification of participants according to ethnic origin (Europids, non-Europids) and by sex.

Approval for the conduct of the ADVANCE study was obtained from the Institutional Ethics Committee of each participating centre and all participants provided written informed consent.

Results

The ADVANCE sample contributing data to the present analyses consisted of 4,010 men and 3,492 women without CVD at baseline and is described in Table 1. To better

Table 1 Baseline profiles of men and women from the ADVANCE population who were free of CVD at the start of the study and of those included in the derivation of the Framingham and UKPDS models

Variable	ADVANCE		Framingham [16]		UKPDS [9][a]	
	Men	Women	Men	Women	Men	Women
n	4,010	3,492	3,969	4,522	2,643	1,894
With diabetes (%)	100	100	6.5	3.8	100	100
Age (years)	66 (6)	66 (6)	48 (11)	49 (11)	51 (9)	53 (9)
Diabetes duration (years)[b]	7 (3–11)	7 (3–11)	NR	NR	0	0
Smoking, n (%)	767 (19)	358 (10)	1,398 (35)	1,548 (34)	898 (34)	474 (25)
Systolic BP (mmHg)	145 (21)	145 (22)	130 (18)	126 (20)	133 (18)	139 (21)
HbA$_{1c}$ (%)	7.5 (1.5)	7.6 (1.6)	NR	NR	6.6 (1.4)	6.9 (1.5)
Total cholesterol (mmol/l)	5.0 (1.1)	5.6 (1.2)	5.5 (1.0)	5.8 (1.1)	5.2 (1.0)	5.7 (1.1)
HDL-cholesterol (mmol/l)	1.2 (0.3)	1.4 (0.4)	1.2 (0.3)	1.5 (0.4)	1.1 (0.2)	1.2 (0.3)
Treated hypertension, n (%)	2,386 (59)	2,462 (70)	402 (10)	532 (12)	NR	NR
Europids, n (%)	2,404 (60)	2,025 (58)	NR	NR	2,151 (81)	1,603 (85)

Unless otherwise indicated, values are mean (standard deviation) or n (%)

[a] Biological variables and systolic BP for the UKPDS cohort are the averages of measurements at 1 and 2 years from baseline

[b] Median (interquartile range)

NR, not reported

understand the extent to which the ADVANCE cohort may be representative of contemporary populations with diabetes, please see our comparison (Electronic supplementary material [ESM] Table 1) of selected baseline characteristics of the overall ADVANCE cohort (11,140 participants) with data from several cross-sectional studies of patients with diabetes from around the world.

Predicted risk of CVD events During the first 4 years of follow-up, 1,003 incident CVD events were recorded, of which 457 were major events, giving a 4-year major CVD event rate (95% CI) of 6.1% (5.6–6.6). The Anderson equation overestimated this rate by 170% with a predicted absolute major CVD event rate of 16.4%; the D'Agostino equation overestimated by 202% (predicted rate 18.4%; Fig. 1). There was no significant difference in the discrimination of these two models (p=0.20). For both equations, the overall discrimination and calibration were poor, with a systematic risk overestimation across all fifths of predicted probabilities (Fig. 2). For the secondary outcome of total cardiovascular events, the magnitude of the risk overestimation was less than for major events, with an improvement of fit within fifths of expected probabilities. However, a systematic deterioration in discrimination was observed, with no significant difference between models (p=0.09). Results for men and women, and for Europids and non-Europids were similar.
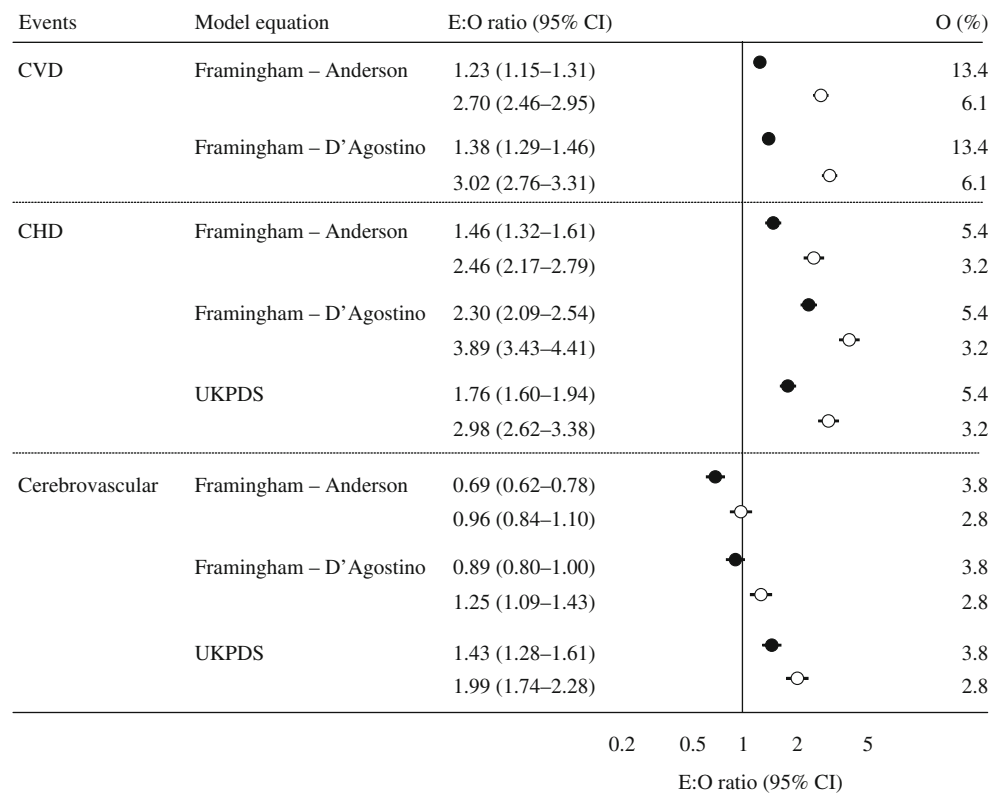
Predicted risk of CHD and cerebrovascular events During follow-up, 407 incident CHD events occurred, of which

241 were major events, giving a 4-year major CHD event rate (95% CI) of 3.2% (2.8–3.6). The overall probability of major CHD was overestimated by the three equations, with the D'Agostino equation performing less well (Fig. 1). Across the fifths of expected probability in the three equations, there was a consistent systematic overestimation of the risks, with a significant lack of fit (Fig. 2). The discriminative power of the three models was low, although the UKPDS discriminated better than the Framingham equations (p<0.02 for UKPDS vs both; Fig. 2). Calibration was improved, but discrimination deteriorated when the outcome any CHD was considered (Figs 1 and 2).

During follow-up, 288 incident cerebrovascular events occurred, of which 207 were major events, giving a 4-year major cerebrovascular event rate (95% CI) of 2.8% (2.4–3.1). The overall probability of a major cerebrovascular event was well estimated by the Anderson equation and overestimated by the UKPDS and D'Agostino equations (Fig. 1). Within fifths of expected risk of stroke, underestimation occurred in the lower fifths and overestimation in the upper fifths, with a consistent significant lack of fit. The three models had low ability to discriminate major cerebrovascular events, but with a significant advantage for the UKPDS equation (p≤0.05 for UKPDS vs both; Fig. 2). When models were assessed for the outcome of any cerebrovascular event, the ability to discriminate deteriorated further, although calibration improved for the UKPDS while worsening for the Framingham equations (Fig. 2).

All these results remained unchanged when analyses were stratified by sex. Stratification by ethnic origin had

Fig. 1 Ratio of 4-year expected (*E*) CVD, CHD and cerebrovascular event rates estimated by the Framingham and UKPDS equations relative to those observed (*O*) in the ADVANCE cohort, with 95% CI. Ratios were estimated using the baseline characteristics for all variables in the 7,502 ADVANCE participants who were free of cardiovascular event at baseline. Point estimates (circles) are for any events (black) and major events (white). Horizontal bars through circles, 95% CI. Vertical line at 1 is the line of perfect agreement between expected and observed events rates

Events	Model equation	E:O ratio (95% CI)	O (%)
CVD	Framingham – Anderson	1.23 (1.15–1.31)	13.4
		2.70 (2.46–2.95)	6.1
	Framingham – D'Agostino	1.38 (1.29–1.46)	13.4
		3.02 (2.76–3.31)	6.1
CHD	Framingham – Anderson	1.46 (1.32–1.61)	5.4
		2.46 (2.17–2.79)	3.2
	Framingham – D'Agostino	2.30 (2.09–2.54)	5.4
		3.89 (3.43–4.41)	3.2
	UKPDS	1.76 (1.60–1.94)	5.4
		2.98 (2.62–3.38)	3.2
Cerebrovascular	Framingham – Anderson	0.69 (0.62–0.78)	3.8
		0.96 (0.84–1.10)	2.8
	Framingham – D'Agostino	0.89 (0.80–1.00)	3.8
		1.25 (1.09–1.43)	2.8
	UKPDS	1.43 (1.28–1.61)	3.8
		1.99 (1.74–2.28)	2.8

0.2   0.5   1   2   5

E:O ratio (95% CI)

variable effects on discrimination and calibration properties. There was a similar overestimation of the risk of CHD events at the subgroup level and within fifths of expected risks among Europids and non-Europids, with a trend toward poorer discrimination statistics in non-Europids. Poorer discrimination was also observed among non-Europids for prediction of cerebrovascular events, while a trend toward better calibration among Europids with the Framingham models and better calibration among non-Europids using the UKPDS models were seen.

Sensitivity analysis In the double placebo cohort (1,882 participants, 52% men, 58% Europids), 253 CVD events (128 major), 108 CHD events (71 major) and 72 cerebrovascular events (51 major) were recorded. The patterns of estimated probabilities of CVD and its components by the three models were broadly comparable to those observed in the overall ADVANCE cohort, with similar magnitudes of risk over-estimation overall, and within sex and ethnic subgroups. For instance, the risk of major CVD events was overestimated by 141% (95% CI 103–187) by the Anderson equation and by 169% (126–220) by the D'Agostino equation. Discrimination was similarly low for the two equations (Anderson AUC [95% CI], 0.61 [0.57–0.66]; D'Agostino 0.62 [0.57–0.66]; *p*=0.69 for the difference). The probability of major CHD events was overestimated by 109% (66–164), 229% (161–315) and 149% (97–214)

respectively by the Anderson, D'Agostino and UKPDS equations. The corresponding AUCs (95% CI) were 0.63 (0.57–0.69), 0.62 (0.56–0.68) and 0.70 (0.65–0.76), with the UKPDS equation performing better than the Framingham ones (*p*<0.004 for UKPDS vs both). The expected risk of major cerebrovascular events was well estimated by the Anderson equation, non-significantly overestimated by the D'Agostino equation and significantly overestimated by the UKPDS equation, with respective expected:observed event ratios of 0.98 (95% CI 0.75–1.29), 1.27 (0.96–1.67) and 1.95 (1.48–2.57). The accompanying AUCs were 0.56 (0.48–0.64), 0.60 (0.52–0.67) and 0.64 (0.56–0.72), with a significant advantage of the UKPDS equation compared with the Anderson (*p*=0.008), but not with the D'Agostino (*p*=0.25) equations.

When the baseline values of blood pressure and HbA$_{1c}$ were replaced in the equations with those at 1-year of follow-up, a modest improvement in discrimination was seen, except for the UKPDS stroke equation, as well as a small attenuation of the magnitude of the estimated risk. However, again, the overall patterns of estimated probabilities of CVD and its components were broadly similar to those from the main analyses. The CIs for the expected/observed event rates ratios at the total population level always overlapped with those from the main analyses, and with one exception, there was always a significant lack of fit within fifths of participants (Fig. 3).
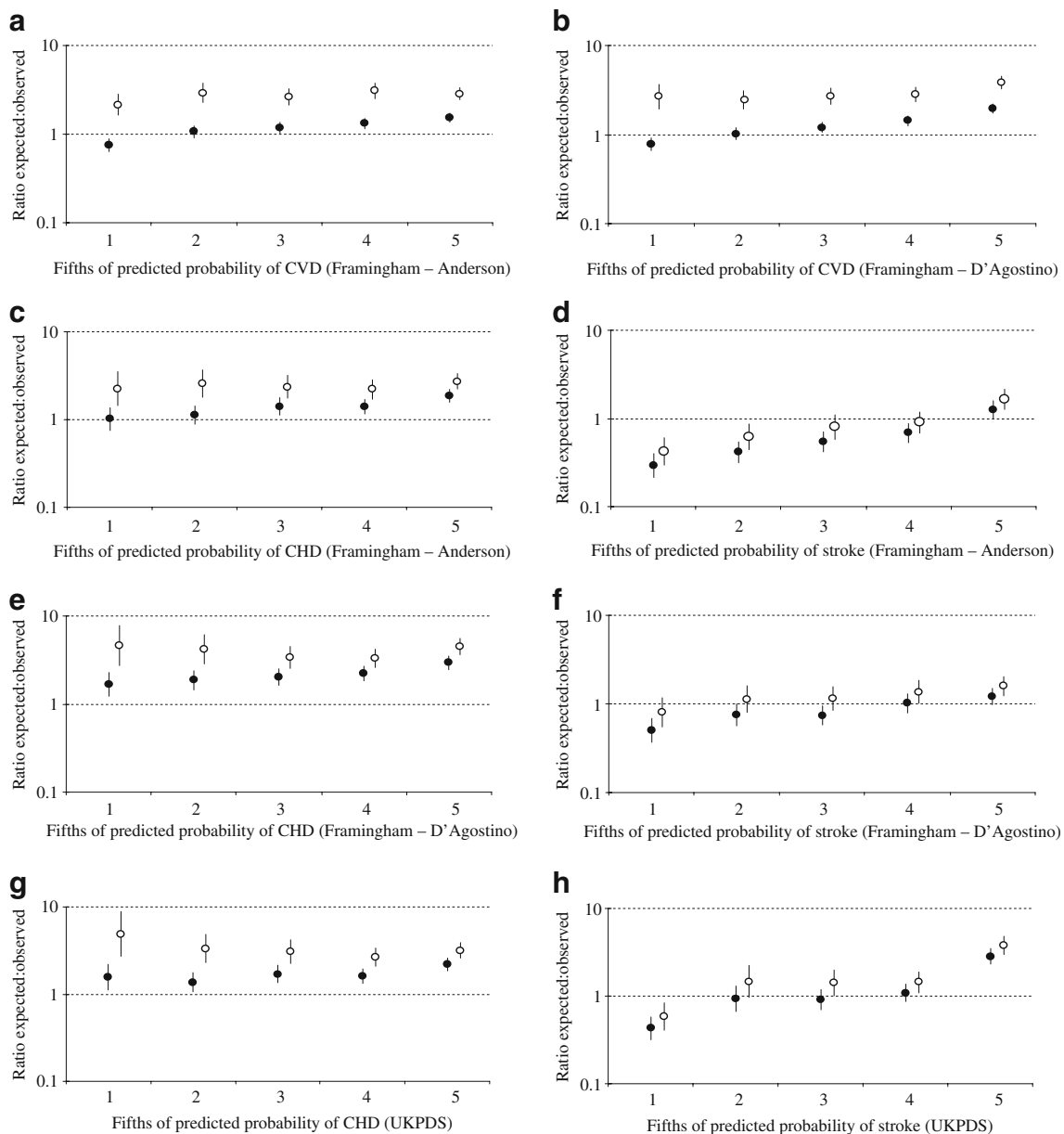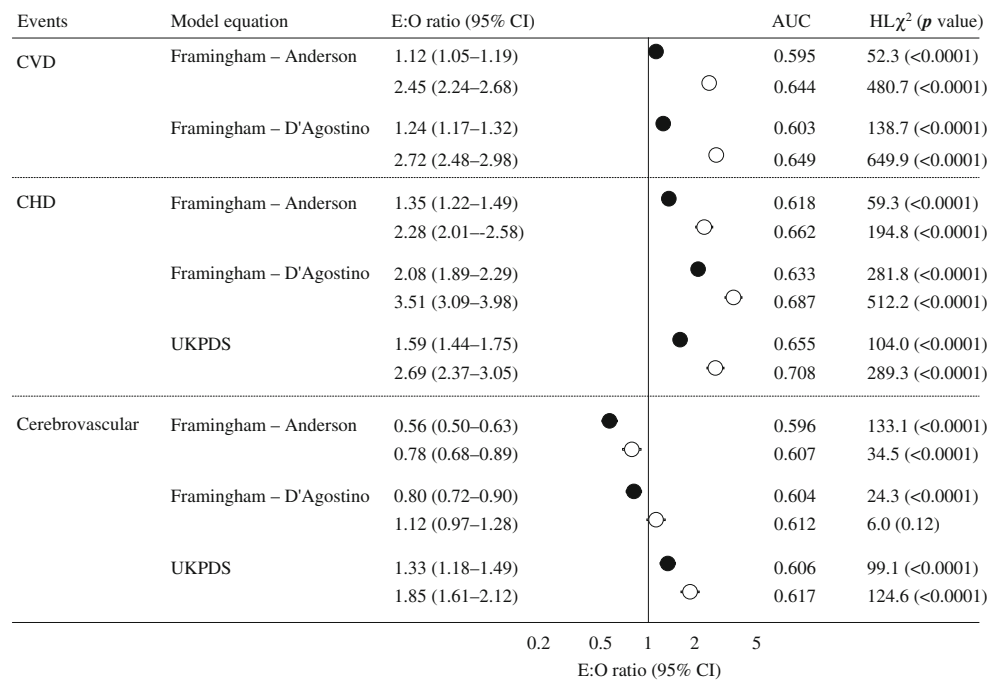
Fig. 2 Ratio of 4-year expected CVD, CHD and cerebrovascular events rates estimated by the Framingham and UKPDS equations relative to those observed in ADVANCE, estimated using the baseline characteristics of the 7,502 ADVANCE participants who were free of cardiovascular event at baseline. Ratios were computed separately within fifths of ADVANCE participants ranked according to the increasing probability of event estimated by each equation. Black circles, any events; white circles, major events; vertical bars through symbols, 95% CI; dotted horizontal line at 1, line of perfect agreement between expected and observed event rates within fifth; dotted horizontal line at 10 was added to assist interpretation. Discrimination (AUC) and Hosmer–Lemeshow calibration χ^2 test (and accompanying p value) are for each model equation (panel) as follows: (**a**) any event 0.579, 87.3 ($p<0.0001$), major event 0.618, 489 ($p<0.0001$); (**b**) any event 0.586, 173.9 ($p<0.0001$), major event 0.625, 621.1 ($p<0.0001$); (**c**) any event 0.611, 71.1 ($p<0.0001$), major event 0.650, 210.2 ($p<0.0001$); (**d**) any event 0.555, 149.0 ($p<0.0001$), major event 0.568, 42.7 ($p<0.0001$); (**e**) any event 0.618, 307.1 ($p<0.0001$), major event 0.666, 518.8 ($p<0.0001$); (**f**) any event 0.567, 32.7 ($p<0.0001$), major event 0.587, 19.9 ($p=0.0004$); (**g**) any event 0.638, 141.1 ($p<0.0001$), major event 0.692, 317.8 ($p<0.0001$); (**h**) any event 0.605, 114.3 ($p<0.0001$), major event 0.618, 138.7 ($p<0.0001$)

Recalibrated models As expected, the recalibrated Anderson Framingham equations accurately predicted the 4-year probability of major events overall, with expected:observed event rates ratios (95% CI) of 1.00 (0.91–1.09) for CVD, 1.00 (0.88–1.14) for CHD and 1.00 (0.87–1.15) for stroke. The equivalents for the D'Agostino equations were 0.90 (0.83–0.99), 1.18 (1.04–1.33) and 0.36 (0.32–0.42) respectively. These represented a massive improvement from the original model for CVD and CHD, and a large underestimation of the risk for stroke. The recalibrated UKPDS CHD and stroke

Events	Model equation	E:O ratio (95% CI)		AUC	HLχ² (p value)
CVD	Framingham – Anderson	1.12 (1.05–1.19)		0.595	52.3 (<0.0001)
		2.45 (2.24–2.68)		0.644	480.7 (<0.0001)
	Framingham – D'Agostino	1.24 (1.17–1.32)		0.603	138.7 (<0.0001)
		2.72 (2.48–2.98)		0.649	649.9 (<0.0001)
CHD	Framingham – Anderson	1.35 (1.22–1.49)		0.618	59.3 (<0.0001)
		2.28 (2.01–2.58)		0.662	194.8 (<0.0001)
	Framingham – D'Agostino	2.08 (1.89–2.29)		0.633	281.8 (<0.0001)
		3.51 (3.09–3.98)		0.687	512.2 (<0.0001)
	UKPDS	1.59 (1.44–1.75)		0.655	104.0 (<0.0001)
		2.69 (2.37–3.05)		0.708	289.3 (<0.0001)
Cerebrovascular	Framingham – Anderson	0.56 (0.50–0.63)		0.596	133.1 (<0.0001)
		0.78 (0.68–0.89)		0.607	34.5 (<0.0001)
	Framingham – D'Agostino	0.80 (0.72–0.90)		0.604	24.3 (<0.0001)
		1.12 (0.97–1.28)		0.612	6.0 (0.12)
	UKPDS	1.33 (1.18–1.49)		0.606	99.1 (<0.0001)
		1.85 (1.61–2.12)		0.617	124.6 (<0.0001)

0.2   0.5   1   2   5
E:O ratio (95% CI)

Fig. 3 Ratio of 4-year expected (*E*) CVD, CHD and cerebrovascular event rates estimated by the Framingham and UKPDS equations, relative to those observed (*O*) in the ADVANCE cohort, with 95% CIs. Ratios were estimated using the 1-year levels of blood pressure and HbA$_{1c}$, and baseline characteristics for all other variables in the 7,502 ADVANCE participants who were free of cardiovascular event at baseline. Point estimates (circles) are for any events (black) and major events (white). Horizontal bars through circles, 95% CI. The vertical line at 1 is the line of perfect agreement between expected and observed events rates. Discrimination (AUC) and Hosmer–Lemeshow calibration χ² test (HLχ²) plus accompanying *p* value are shown for each model equation

functions systematically overestimated the rates of major CHD and stroke. However, the magnitude of such overestimation was less than that observed using the original equations (overall bias, original vs recalibrated models 197.7% vs 76.3% for CHD, 138.72% vs 99.8% for stroke; Fig. 4).

Discussion

These analyses show that, with few exceptions, the Framingham general cardiovascular risk equations and the UKPDS diabetes-specific coronary and stroke risk equations were poor at predicting the risk of CVD events or its constituents in the ADVANCE cohort. These equations were more likely to overestimate risk and even when the overall risk of the cohort was well estimated, they had only limited power to discriminate between high- and low-risk individuals. Discrimination was better and calibration poorer when prediction was restricted to 'hard' major events. Similar model performance was observed in men and women, and in Europids and non-Europids, and also when analyses were restricted to the cohort allocated to control, thus accounting for any potential effects of the randomised interventions in the ADVANCE study. Adequate improvement in fit through recalibration was achieved for major

CVD and CHD in general, but only for stroke with the UKPDS equation.

A number of CVD risk prediction scores, including those derived from Framingham and UKPDS, have been previously evaluated in independent populations of diabetic patients [25–28]. A minority of these studies have reported on measures of model discrimination and calibration. Discriminatory capability, mainly assessed by AUC under the ROC curve, was modest to acceptable and varied substantially according to the type of CVD outcome, as we found with the ADVANCE cohort. The overall predicted risk of CVD in these studies ranged from under- to overestimation, but calibration within subgroups, where assessed using the Hosmer–Lemeshow statistic, was consistently poor [25, 26].

The poor performance of the Framingham and UKPDS risk equations in the ADVANCE cohort can be explained, at least in part, by differences in the baseline profiles of the three populations. Unlike the Framingham study and UKPDS, the ADVANCE study included a broad cross-section of relatively older participants from many countries and ethnic backgrounds. Such differences are highly likely to influence model calibration. Moreover, and as indicated by the recalibration process in this analysis, accounting for these differences improved model performance, at least in part. The properties of the statistical approaches used for
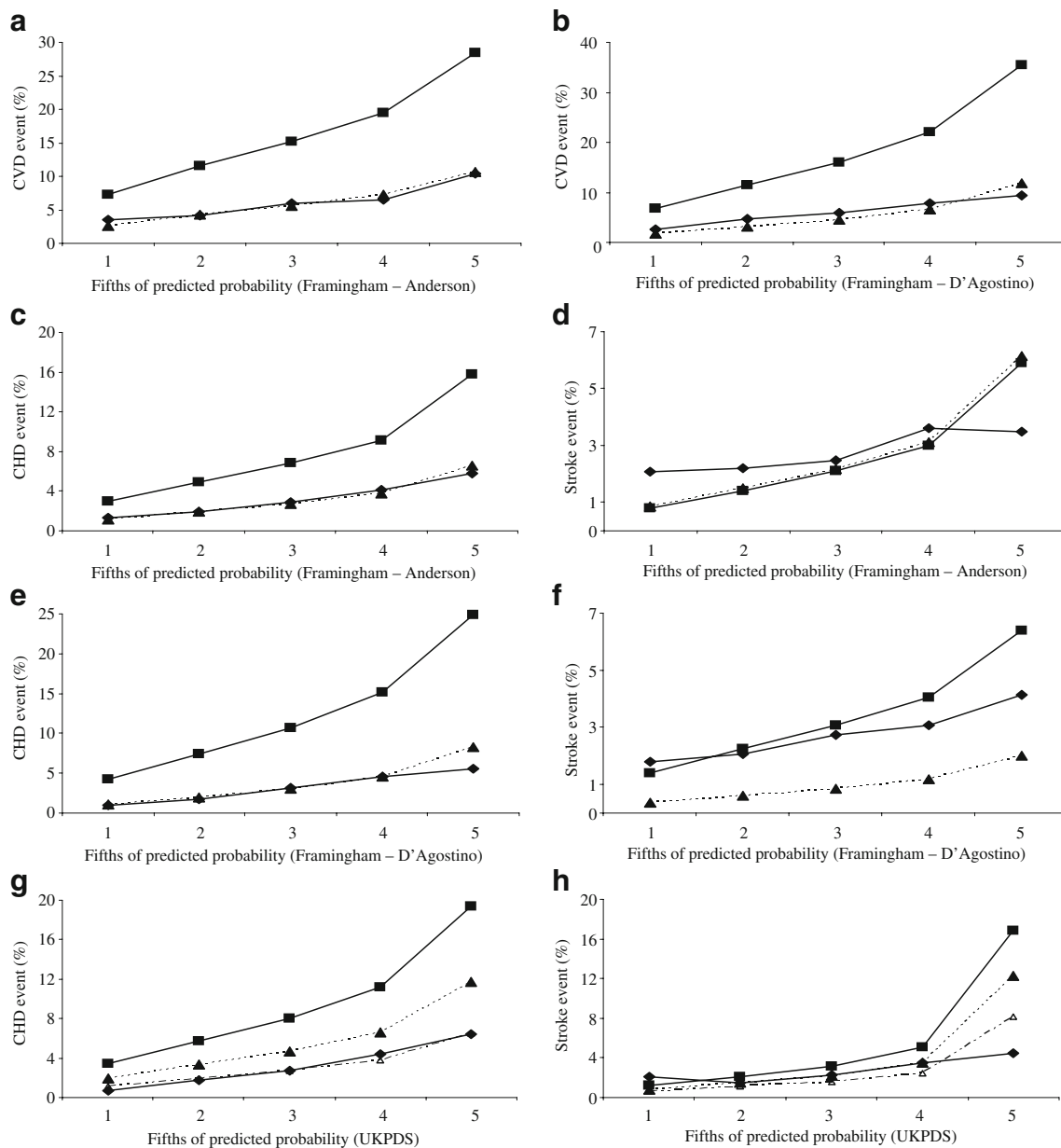
Fig. 4 Expected and observed major CVD, CHD and cerebrovascular event rates by fifths of predicted probability according to the Framingham and UKPDS equations for the original and the recalibrated models. The observed (diamonds) and expected event rates from the original (squares) and recalibrated (triangles) models were computed separately within fifths of ADVANCE participants, ranked according to the increasing probability of event estimated by each equation. Interpolation lines have been added to relate estimates from consecutive fifths and assist interpretation. The overall bias in estimated event rates by each equation (panel) before and after recalibration was: (**a**) 170% and 0%, (**b**) 202% and −10%, (**c**) 236.2% and 0.1%, (**d**) −3.7% and −0.1%, (**e**) 289% and 17.6%, (**f**) 24.5% and −63.6%, (**g**) original equation only, 198% and (**h**) original equation only, 99%. For the UKPDS equations (**g, h**), one recalibration (black triangles) was based on the ADVANCE-specific level of predictors at baseline; another recalibration (white triangles) was based on the ratio of observed:expected event rates [23]. The overall bias in the estimated event rates according to the two recalibrated models was 76% and 0% for CHD (**g**), and 58% and 0% for stroke (**h**), respectively

assessing model performance could also contribute to some of the observed discrepancies. For example, a disadvantage of the Hosmer–Lemeshow test is that the value of the statistic is sensitive to: (1) the choice of cut-off points that define the subgroups; (2) the number of such subgroups; and (3) the sample size. Using small samples will result in an apparently good fit; using large samples will result in an apparently poor fit [29].

It is not possible to obtain perfect calibration and discrimination. Since discrimination cannot be modified by subsequent recalibration or by modifying treatment thresholds, it is often regarded as the most important

property of a risk prediction score. The discrimination properties of the models when applied to the ADVANCE cohort were modest to acceptable. This was probably due to the fact that some important predictors in the ADVANCE cohort were not included in the UKPDS and Framingham equations. Such predictors may include variables relating to preventative therapies. For instance, ADVANCE participants were an intensively treated group, with over 75% on blood pressure-lowering and 50% on lipid-modifying medications at study completion, respectively [14]. These treatments are likely to affect the association between baseline levels of predictors and outcomes during follow-up. In addition, ADVANCE participants were older than the UKPDS and Framingham populations. Many epidemiological studies have indicated that the relative risks of CVD associated with major risk factors attenuate with ageing [30–32]. The parametric form of the UKPDS models and the assumptions underlying the integration of some predictors in the models could also have significantly influenced the discrimination properties of these models.

A risk prediction model provides accurate estimates of the outcomes only when it is used in a population with characteristics similar to those of the population from which the model was derived. However, appropriate data and resources will not always be available to develop new models specific for each setting. Strategies to modify existing models, including model recalibration, have been proposed with a view to improving the performance of risk prediction models in different populations. These approaches have been used, for example, to adapt the Framingham cardiovascular risk equations to different populations [33]. In the only available study of such an approach, which was conducted in people with diabetes, no improvement was found in the discrimination property of the Framingham Anderson model, after refitting the equation to the data from Cardiff [34], but calibration performance was not reported. In the ADVANCE cohort, recalibrated models showed greatly improved calibration, suggesting that recalibration could be used to improve the performance of the Framingham and UKPDS models in people with diabetes. For instance, if validated in another contemporary cohort with diabetes, the recalibration variables derived from the ADVANCE sample could be used in other settings where follow-up data are not available to adjust the risk estimates from the Framingham and UKPDS equations in people with diabetes. It should, however, be noted that improvement in model performance after recalibration was not always optimal, suggesting that the method may not be valid in all circumstances. For instance, an unacceptable underestimation of the risk of stroke was observed with the Framingham D'Agostino equation and a significant residual risk overestimation with the two UKPDS equations after recalibration.

The value of the present analysis depends largely on the extent to which the ADVANCE participants are typical of current populations of patients with type 2 diabetes managed in the community. Comparison with available representative cross-sectional data suggests a broad similarity with populations described in a number of settings, although only 1% of ADVANCE participants were on insulin treatment despite an average duration of diabetes of 7 years. However, it must be acknowledged that, by the nature of trial inclusion criteria, ADVANCE participants were more likely to be intermediate- or high-risk patients [14, 15]. On this basis, one would have expected the Framingham and UKPDS equations to underestimate rather than overestimate risk when used on the ADVANCE sample. The overestimation observed is likely to be even greater when using these equations in a more general population with diabetes (and presumably at lower risk) as suggested elsewhere [27, 28]. The predictive accuracy of the models in the ADVANCE study could only be tested over a follow-up period of 4 years. Validating these tools over a longer follow-up period may alter the findings. The choice of a 4-year follow-up was motivated by the need to separately assess the performance of the models in the 'double control' group of the ADVANCE study. The mean duration of follow-up of this group was determined by the length of follow-up of the blood pressure arm of the ADVANCE study and was 4.3 years. Model recalibration based on the observed/expected event rates ratio is a rigid approach based on the inherent assumption that an observed/expected ratio will remain constant across various subgroups within a given population. As shown in this analysis, the magnitude and direction of departures of the predicted event rates/number by risk tools from those observed can vary substantially across populations [35]. Several recalibration methods with variable levels of complexity have been suggested for updating survival models for new settings, with no indication that the more complex methods provide the best results [24, 34, 36].

To date, this study is the largest, most comprehensive validation study of cardiovascular risk prediction in a diverse and contemporary population with type 2 diabetes. Unlike most previous validation studies, we assessed the effects of recalibration on model performance. Providing precision on risk estimates will facilitate reliable comparison of our findings with available and subsequent studies. The inclusion in the ADVANCE cohort of participants from many countries has enabled us to simultaneously assess key aspects of the generalisability of the UKPDS and Framingham equations, including their reproducibility and transportability (historical, geographical and methodological) [37]. There are several important clinical implications arising from the inaccuracy in absolute risk prediction by the Framingham and UKPDS risk equations. Failure to discriminate risk well is a serious

issue, since risk prediction scores are primarily used to ascertain relative levels of risk and prioritise treatment of individuals. Absolute risk equations are also helpful in communicating information on prognosis to patients and their care-providers, as well as in estimating the balance of potential benefits and risks of individual preventive treatments. For these reasons, precise estimation of risks is important. Both the Framingham and UKPDS equations substantially overestimated the risks in the ADVANCE cohort. The magnitude of this overestimation was attenuated through recalibration; however, this improvement was incomplete. As well as recalibration, the creation of new risk prediction tools with improved predictive accuracy would help to tailor preventative strategies even more reliably in people with diabetes, particularly among those already receiving multiple contemporary treatments.

References

1. Gordon T, Kannel WB (1982) Multiple risk functions for predicting coronary heart disease: the concept, accuracy, and application. Am Heart J 103:1031–1039
2. International Task Force for Prevention of Coronary Heart Disease, International Atherosclerosis Society (2003) Pocket guide to prevention of coronary heart disease. Born Bruckmeier, Munster
3. Expert Panel on Detection, Evaluation and Treatment of High Blood Cholesterol in Adults (2002) Third report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. Circulation 106:3143–3421
4. Chobanian AV, Bakris GL, Black HR et al (2003) The seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. JAMA 289:2560–2572
5. Haffner SM, Lehto S, Ronnemaa T, Pyorala K, Laakso M (1998) Mortality from coronary heart disease in subjects with type 2 diabetes and in nondiabetic subjects with and without prior myocardial infarction. N Engl J Med 339:229–234
6. Asia Pacific Cohort Studies Collaboration (2006) Coronary risk prediction for those with and without diabetes. Eur J Cardiovasc Prev Rehabil 13:30–36
7. Woodward M, Brindle P, Tunstall-Pedoe H (2007) Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). Heart 93:172–176
8. Anderson KM, Odell PM, Wilson PW, Kannel WB (1991) Cardiovascular disease risk profiles. Am Heart J 121:293–298
9. Stevens RJ, Kothari V, Adler AI, Stratton IM (2001) The UKPDS risk engine: a model for the risk of coronary heart disease in type II diabetes (UKPDS 56). Clin Sci (Lond) 101:671–679
10. Anderson KM, Wilson PW, Odell PM, Kannel WB (1991) An updated coronary risk profile. A statement for health professionals. Circulation 83:356–362
11. Kothari V, Stevens RJ, Adler AI et al (2002) UKPDS 60: risk of stroke in type 2 diabetes estimated by the UK Prospective Diabetes Study risk engine. Stroke 33:1776–1781
12. Donnan PT, Donnelly L, New JP, Morris AD (2006) Derivation and validation of a prediction score for major coronary heart disease events in a U.K. type 2 diabetic population. Diabetes Care 29:1231–1236
13. Stevens RJ, Coleman RL, Adler AI, Stratton IM, Matthews DR, Holman RR (2004) Risk factors for myocardial infarction case fatality and stroke case fatality in type 2 diabetes: UKPDS 66. Diabetes Care 27:201–207
14. Patel A, MacMahon S, Chalmers J et al (2007) Effects of a fixed combination of perindopril and indapamide on macrovascular and microvascular outcomes in patients with type 2 diabetes mellitus (the ADVANCE trial): a randomised controlled trial. Lancet 370:829–840
15. Patel A, MacMahon S, Chalmers J et al (2008) Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. N Engl J Med 358:2560–2572
16. D'Agostino RB Sr, Vasan RS, Pencina MJ et al (2008) General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 117:743–753
17. DeLong ER, DeLong DM, Clarke-Pearson DL (1988) Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 44:837–845
18. Rockhill B, Spiegelman D, Byrne C, Hunter DJ, Colditz GA (2001) Validation of the Gail et al. model of breast cancer risk prediction and implications for chemoprevention. J Natl Cancer Inst 93:358–366
19. Asia Pacific Cohort Studies Collaboration (2007) Cardiovascular risk prediction tools for populations in Asia. J Epidemiol Community Health 61:115–121
20. Hosmer DW, Lemeshow S (1989) Applied logistic regression. Wiley, New York
21. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB (1998) Prediction of coronary heart disease using risk factor categories. Circulation 97:1837–1847
22. Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB (1991) Probability of stroke: a risk profile from the Framingham Study. Stroke 22:312–318
23. Brindle P, Emberson J, Lampe F et al (2003) Predictive accuracy of the Framingham coronary risk score in British men: prospective cohort study. BMJ 327:1267
24. van Houwelingen HC (2000) Validation, calibration, revision and combination of prognostic survival models. Stat Med 19:3401–3415
25. Chamnan P, Simmons RK, Sharp SJ, Griffin SJ, Wareham NJ (2009) Cardiovascular risk assessment scores for people with diabetes: a systematic review. Diabetologia 52:2001–2014
26. Davis WA, Colagiuri S, Davis TM (2009) Comparison of the Framingham and United Kingdom Prospective Diabetes Study cardiovascular risk equations in Australian patients with type 2 diabetes from the Fremantle Diabetes Study. Med J Aust 190:180–184
27. Simmons RK, Coleman RL, Price HC et al (2009) Performance of the UK Prospective Diabetes Study risk engine and the Framingham risk equations in estimating cardiovascular disease in the EPIC-Norfolk Cohort. Diabetes Care 32:708–713
28. van der Heijden AA, Ortegon MM, Niessen LW, Nijpels G, Dekker JM (2009) Prediction of coronary heart disease risk in a general, pre-diabetic, and diabetic population during 10 years of follow-up: accuracy of the Framingham, SCORE, and UKPDS risk functions: The Hoorn Study. Diabetes Care 32:2094–2098

29. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S (1997) A comparison of goodness-of-fit tests for the logistic regression model. Stat Med 16:965–980

30. Kannel WB, D'Agostino RB (1995) The importance of cardiovascular risk factors in the elderly. Am J Geriatr Cardiol 4:10–23

31. Kannel WB (2002) Coronary heart disease risk factors in the elderly. Am J Geriatr Cardiol 11:101–107

32. Moller CS, Zethelius B, Sundstrom J, Lind L (2006) Impact of follow-up time and re-measurement of the electrocardiogram and conventional cardiovascular risk factors on their predictive value for myocardial infarction. J Intern Med 260:22–30

33. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P (2001) Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. JAMA 286:180–187

34. McEwan P, Williams JE, Griffiths JD et al (2004) Evaluating the performance of the Framingham risk equations in a population with diabetes. Diabet Med 21:318–323

35. Brindle P, Beswick A, Fahey T, Ebrahim S (2006) Accuracy and impact of risk assessment in the primary prevention of cardiovascular disease: a systematic review. Heart 92:1752–1759

36. Steyerberg EW (2009) Clinical prediction models: a practical approach to development, validation, and updating. Springer, Rotterdam

37. Justice AC, Covinsky KE, Berlin JA (1999) Assessing the generalizability of prognostic information. Ann Intern Med 130:515–524