Derivation and Validation of a Prediction Score for Major Coronary Heart Disease Events in a U.K. Type 2 Diabetic Population

PETER T. DONNAN, PHD[1]
LOUISE DONNELLY, BSC[2]
JOHN P. NEW, FRCP[3]
ANDREW D. MORRIS, MD[2,4]

OBJECTIVE — To derive and validate an absolute risk algorithm for major coronary heart disease (CHD) events in the U.K. population with type 2 diabetes.

RESEARCH DESIGN AND METHODS — A population cohort with type 2 diabetes was constructed in Tayside, Scotland, U.K., and longitudinally followed-up to June 2004. Participants were all people with type 2 diabetes registered with general practices and the Diabetes Audit and Research in Tayside, Scotland, database (97% sensitive) with no previous CHD event and with complete measurements (n = 4,569). The main outcome measure was risk of CHD defined as fatal or nonfatal myocardial infarction or CHD death, derived from the Weibull accelerated failure-time model. Validation of the algorithm was performed on an independent dataset from Salford, England, U.K.

RESULTS — There were a total of 243 subjects (5.3%) with a fatal or nonfatal myocardial infarction or CHD death over the follow-up period from 1 January 1995 to 30 June 2004 (maximum follow-up 9.5 years). The final Weibull model included the significant predictors of age at diagnosis, duration of diabetes, HbA_{1c}, smoking (current, past, never), sex, systolic blood pressure, treated hypertension, total cholesterol, and height. Assessment of discrimination and calibration in the Salford validation dataset demonstrated a good fit (c = 0.71 [95% CI 0.63–0.79]).

CONCLUSIONS — This study provides the first validated, population-derived model for prediction of absolute risk of CHD in people with type 2 diabetes. It provides a useful additional decision aid for the clinician treating type 2 diabetes by indicating appropriate early action to decrease the risk of adverse outcomes.

Diabetes Care **29:1231–1236, 2006**

Patients with type 2 diabetes have a significantly increased risk of developing coronary heart disease (CHD), in part due to the clustering of numerous risk factors including hypertension and dyslipidemia (1,2). In diabetic individuals with no previous history, the risk of developing an acute myocardial infarction (AMI) may be equivalent to that of a nondiabetic individual with a previous event (3). Following AMI, diabetic patients are more likely to develop cardiac failure and cardiogenic shock, resulting in a greater short- and long-term mortality (4). The primary and secondary prevention of CHD in patients with diabetes is therefore key to successful diabetes management. There is convincing evidence that blood pressure– and cholesterol-lowering treatment is effective for both the primary and secondary prevention of CHD (1,5,6), and recent large primary prevention studies including the Heart Protection Study have demonstrated that treatment with statins was also effective and safe for diabetic patients at high risk of vascular events (7). In the Hypertension Optimal Treatment study (8), aspirin was shown to lower risk of cardiovascular events by lowering blood pressure, and aspirin therapy has also been endorsed by the American Diabetes Association for high-risk patients (9).

A recent review (10) suggested that separate management guidelines for blood pressure and blood cholesterol should be replaced by integrated cardiovascular risk management and the routine use of absolute cardiovascular risk prediction scores. However, the use of algorithms that estimate the risk of cardiovascular disease in the important subgroup of people with diabetes is controversial (11). Most risk equations are based on the Framingham Study, which only included 337 people with diabetes in their cohort of ~5,000 (12,13). An alternative algorithm is based on the Prospective Cardiovascular Munster Study (14), but this was derived from a working population of middle-aged men, which not only does not permit estimation for women but also suffers from the "healthy worker effect." The applicability of these algorithms to certain subgroups is contentious; specifically, they are thought to underestimate the risk of CHD in people with type 2 diabetes (15). In particular, they take no account of diabetes-specific risk factors such as glycemic control and duration of diabetes.

Most U.K. and European recommendations on the primary prevention of CHD in the general population (16) have based a decision on treatment with lipid-lowering therapy upon an estimate of the

From the [1]Tayside Centre for General Practice, Health Informatics Centre, Community Health Sciences, University of Dundee, Dundee, U.K.; the [2]Medicines Monitoring Unit, Health Informatics Centre, Community Health Sciences, University of Dundee, Dundee, U.K.; the [3]Department of Diabetes, Hope Hospital, Salford, U.K.; and the [4]Division of Medicine and Therapeutics, University of Dundee, Dundee, U.K.

Address correspondence and reprint requests to Dr. Peter T. Donnan, University of Dundee, Tayside Centre for General Practice, Health Informatics Centre, Community Health Sciences, Mckenzie Building, Dundee DD2 4BF, U.K. E-mail: p.t.donnan@chs.dundee.ac.uk.

Received for publication 7 October 2005 and accepted in revised form 21 February 2006.

J.P.N. has received honoraria and grant/research support from Pfizer.

Abbreviations: AMI, acute myocardial infarction; CHD, coronary heart disease; DARTS, Diabetes Audit and Research in Tayside, Scotland; DIS, Salford Diabetes Information System; SBP, systolic blood pressure; UKPDS, U.K. Prospective Diabetes Study.

A table elsewhere in this issue shows conventional and Système International (SI) units and conversion factors for many substances.

DOI: 10.2337/dc05-1911

absolute risk from Framingham data (13). For secondary prevention, it is recommended that total cholesterol concentrations should be <5.0 mmol/l in individuals aged <75 years with documented vascular disease (16–18).

In light of the excess CHD risk in people with diabetes, there is a need for a robust algorithm that reliably estimates the risk of CHD for individuals in the population with type 2 diabetes. A previous risk equation derived from the U.K. Prospective Diabetes Study (UKPDS) attempted to provide this (18). However, this risk equation was derived from a randomized, controlled trial sample with age restrictions and exclusions for comorbidity. The authors also acknowledged selection bias and did not recommend its use for predictions of <4 years or for people aged >65 years (19), thereby excluding ~60% of patients in a typical clinic. To the best of our knowledge, a risk equation derived from an unselected population with type 2 diabetes has not yet been developed. The aim of this study was therefore to develop a robust algorithm for clinical use in predicting CHD in subjects with type 2 diabetes in a U.K. population and validate the algorithm in an independent, population-based dataset.

RESEARCH DESIGN AND METHODS

The study population consisted of all subjects with a diagnosis of type 2 diabetes residing in Tayside, Scotland (population ~400,000), and registered with a Tayside general practice from 1 January 1995 to 30 June 2004. Ascertainment of type 2 diabetes was achieved through the Diabetes Audit and Research in Tayside, Scotland (DARTS), database, which has been described in detail elsewhere (20). In brief, record linkage of multiple data sources has been used to create a web-based regional diabetes information system for all residents in Tayside. The sensitivity and positive predictive value of DARTS are both 97% for ascertainment of diagnosis of diabetes (20). For the purpose of this study, subjects with type 2 diabetes were defined as those of any age treated with diet or oral hypoglycemic agents or those aged >35 years at diagnosis. Approximately 15% of those with type 2 diabetes in the database are treated with insulin, while <1% were of nonwhite origin.

Baseline measurements

For each individual identified, a baseline time point was constructed such that all the risk factors under consideration were present within a period of 3 months. The potential risk factors were duration of diabetes (or time from diagnosis of type 2 diabetes to baseline); age at diagnosis; sex; smoking history (current, past, never); total, LDL, and HDL cholesterol; triglycerides (fasting and nonfasting); HbA_{1c} (A1C) (Diabetes Control and Complications Trial standardized, nondiabetic range 4.5–6.2%; Bio-Rad); systolic (SBP) and diastolic blood pressure (mmHg); BMI (kg/m^2); and height (m). Both duration of known diabetes and age at diagnosis were considered, as a previous study suggested that both variables were independently significant predictors rather than age at baseline alone (19).

Exclusions

Those people with incomplete records for one or more potential risk factors were excluded. In addition, people who had experienced a previous cardiovascular disease event (ICD-10: I20, 21, 22, 23, 50, I60-I69) were also excluded.

Outcome

The main outcome for this study was the first presentation with major CHD events defined as fatal or nonfatal AMI (ICD-10 I21, 22, 23) or CHD death (ICD-10 I20–25, 46, 50). Events of hospitalization for AMI were obtained from the Tayside part of the Scottish morbidity register that contains all records of hospitalization and is held in the medicines monitoring unit (21). Myocardial infarction events that did not result in hospitalization were obtained from the DARTS database based on information derived from all general practices in the region (20). Hence, ascertainment of events was close to 100%. All cause of death information was obtained from the general registry office for Scotland. All study events were validated anonymously by a clinical specialist (A.D.M.). Ethics and Caldicott Guardian approval was obtained for this purely electronic study using anonymized data according to the standard operating procedures of the Health Informatics Center (21).

Validation dataset

The Salford Diabetes Information System (DIS) is an independent, population-based diabetes dataset established in the Salford district (population = 215,000) of England, U.K., in 1992 to support an integrated diabetes care program. The DIS has been previously described in detail (22). Briefly, the DIS represents a continuously updated diabetes health care record that prompts structured diabetes care using an automated recall system. Key processes of diabetes care (e.g., weight, blood pressure, and glycemic control measurement; retinal screening; and foot examination) are prompted at an annual structured preventative care review using a standard clinic proforma. These records are updated and verified during the annual review. Outcome data, such as myocardial infarction, are validated yearly by linking the details recorded on DIS with outpatient and inpatient episodes coded through hospital episode statistics (23).

Statistical methods

The characteristics of the study population were summarized by means and SDs for continuous measurements and as percentages for categorical factors. To assess the representativeness of the study population compared with all individuals with type 2 diabetes in Tayside, those excluded because of incomplete measurements were compared with those included in the study population by calculating differences for each factor with 95% CIs. Time to CHD was modeled using both the proportional and nonproportional hazards Weibull accelerated failure-time model (24), as this parametric approach has been found to be more flexible in predicting risk for the Framingham Heart Study (13). Generally, in this model a negative value for the regression coefficient indicates greater risk associated with that factor.

Factors for potential inclusion in the model were initially considered with univariate significance of $P < 0.20$ and/or judged to be of clinical importance. In fitting the regression model, the main factors used in the Framingham prediction equations, namely age; sex; smoking; SBP; and total, LDL, and HDL cholesterol, were simultaneously assessed in the model (13). In addition, the potential predictors of duration of diabetes (or time from diagnosis), age at diagnosis, A1C, BMI, triglycerides, and height were also added, if P was <0.05 or of clinical importance. Tests for interactions between risk factors were performed at the $P < 0.05$ level of significance. Finally, rejected variables were added again to the final model to confirm that they were not statistically significant, clinically important, or potential confounders. Treatment for hypertension was recorded in the DARTS

Table 1—*Characteristics of the study population (n = 4,569) and those excluded with incomplete data (n = 9,170)*

Variable	Complete	Incomplete
n	4,569	9,170
Age at diagnosis (years)	54.7 ± 12.2	65.4 ± 12.8 (*n* = 8,698)
Duration of diabetes (years)	4.7 ± 6.1	NA*
Age at baseline (years)	59.5 ± 12.1	NA*
Sex		
Men	52.6	51.2 (*n* = 9,170)
Women	47.4	48.8
Current smoker	23.5	17.5 (*n* = 7165)
Ex-smoker	20.4	33.7
Nonsmoker	56.1	48.7
Treated hypertension	61.9	53.7
A1C (%)	8.0 ± 1.9	7.9 ± 1.8 (*n* = 411)
SBP (mmHg)	144 ± 21	145 ± 17 (*n* = 7,585)
DBP (mmHg)	82 ± 11	80 ± 8.7 (*n* = 7,583)
Total cholesterol (mmol/l)	5.4 ± 1.1	5.5 ± 1.3 (*n* = 1,092)
HDL cholesterol (mmol/l)	1.2 ± 0.4	1.2 ± 0.5 (*n* = 950)
Log (total/HDL)	1.52 ± 0.34	1.55 ± 0.37 (*n* = 902)
Triglycerides (mmol/l)	2.8 ± 2.1	2.92 ± 2.0
BMI (kg/m^2)	30.7 ± 6.4	29.3 ± 5.8 (*n* = 6,790)
Height (m)	1.67 ± 0.1	1.66 ± 0.1

Data are means ± SD or percent. *No baseline so not possible to estimate. DBP, diastolic blood pressure.

database, and this was also assessed as a risk factor.

The assumption of proportional hazards was assessed by plotting the log-negative-log plot for individual factors, fitting time-dependent variables, and by trend tests using the Schonfeld residuals (25). The estimated probabilities of CHD were obtained from the final model with 95% CIs using the delta method (26).

Model performance

Performance of the algorithm obtained from the Tayside dataset was tested on the independent Salford dataset. First, the overall discrimination ability was assessed for the Tayside function on Tayside data and, second, using the Tayside model on the Salford data. Discrimination was assessed using the *c* statistic, which is an estimate of the probability of assigning a higher risk to those who develop CHD in 5 years compared with those who do not (26). Regression coefficients were compared for the derivation model applied to the Salford dataset using z-tests. Finally, the Gronnesby and Borgan test was carried out by adding deciles of predicted risk to the model applied to the validation dataset giving a partial likelihood test for calibration (27). All analyses were implemented in SAS Version 9 (SAS Institute, Cary, NC).

RESULTS — A total of 4,569 individuals with type 2 diabetes were included, after exclusion of those with previous cardiovascular events and/or with incomplete measurements. Table 1 shows the characteristics of the study population. The study population consisted of 52.6% men and 47.4% women, with a mean ± SD age at baseline of 59.5 ± 12.1 years. Those individuals with type 2 diabetes who were excluded from the study were on average older (mean age 65 vs. 55 years), but their clinical measurements were very similar to the study population. For example, in Table 1, mean A1C is 8.0 vs. 7.9%, SBP 144 vs. 145 mmHg, and total cholesterol 5.4 vs. 5.5 mmol/l.

There were a total of 18,831 person-years of follow-up in the study. A total of 243 subjects (5.3%) suffered a fatal or nonfatal AMI or CHD death over the maximum follow-up period from 1 January 1995 to 30 June 2004 (median 4.1 years [range 9.5]). In the standard Weibull model, longer duration of diabetes was highly significantly associated with increased risk of CHD, along with and independently of age at diagnosis (Table 2). Men had significantly greater risk of CHD compared with women. High cholesterol and high SBP were also strongly associated with increased risk. There was a significant interaction between SBP and treated hypertension with higher risk of CHD in untreated patients with high SBP. Current smoking was significantly associated with increased risk, while ex-smokers had a higher risk compared with nonsmokers but lower than current smokers. A1C as a log-transformed factor was associated with increased risk. Shorter stature was univariately associated with increased risk, while BMI, triglycerides, and HDL cholesterol showed no significant univariate association.

All variables, where univariately $P < 0.2$, were considered for entry in a multivariate model. The test for slope of the time-varying coefficient was significant for A1C, indicating nonproportional hazards for this factor. The nonproportional hazards Weibull model was applied with

Table 2—*Final Weibull model for prediction of risk of CHD*

Factor	Coefficient (95% CI)	*P* value
Intercept	11.262 (8.943–13.582)	<0.0001
Log (duration of diabetes)	−0.287 (−0.362 to −0.211)	<0.0001
Age at diagnosis	−0.026 (−0.034 to −0.018)	<0.0001
Total cholesterol	−0.149 (−0.216 to −0.081)	<0.0001
Smoking status		
Nonsmoker	—	
Ex-smoker	0.011 (−0.187 to 0.209)	0.915
Current smoker	−0.268 (−0.450 to −0.086)	0.004
Men versus women	−0.308 (−0.522 to −0.095)	0.005
Log (A1C)	0.438 (0.073–0.802)	0.019
Log (A1C) × follow-up (≤5 vs. >5 years)	−0.712 (−0.812 to −0.611)	<0.0001
SBP	−0.010 (−0.017 to −0.003)	0.004
Treated hypertension (yes vs. no)	−1.292 (−2.483 to −0.100)	0.034
SBP × treated hypertension	0.009 (0.001–0.018)	0.021
Height	1.241 (0.143–2.339)	0.027
σ	0.587 (0.527–0.655)	

one additional parameter, but the improvement in fit over the standard Weibull model was not significant (likelihood ratio test, $P = 0.054$). The hazard plots for A1C were approximately proportional in the first 5 years. Consequently, a final model with interactions of a binary indicator for <5 years of follow-up (yes/no) with log-transformed A1C was added, which significantly improved the goodness of fit (Table 2). As an example of the calculation of risk, consider a male current smoker, with 6 years' duration of diabetes and diagnosis at the age of 59 years, SBP 160 mmHg with no treatment, total cholesterol 5.8 mmol/l, A1C 8%, and height 1.7 m, giving a 5-year risk of a CHD event of 54% (95% CI 42–67).

The discriminatory power for the prediction algorithm on the Tayside data were $c = 0.71$ (95% CI 0.63–0.79). When the Tayside algorithm was applied to the Salford data, discrimination was still good with $c = 0.69$ (0.58–0.78). When the differences between the regression coefficients for the Tayside algorithm and the regression coefficients when applied to the Salford data were tested, there was only one factor that showed a difference. The difference for the log of duration of diabetes reached statistical significance ($P = 0.04$), although if significance levels were adjusted for multiple testing this would not be considered significant. When nine design variables for the deciles of predicted risk based on the Tayside algorithm were added to the model when applied to the validation dataset, the partial likelihood test gave $G = 7.94$ ($\chi^2_9 = 7.94$, $P = 0.540$). The actual CHD events in Salford and predicted over 5 years by decile of predicted risk are presented in Fig. 1. These demonstrate a good fit to the validation dataset.

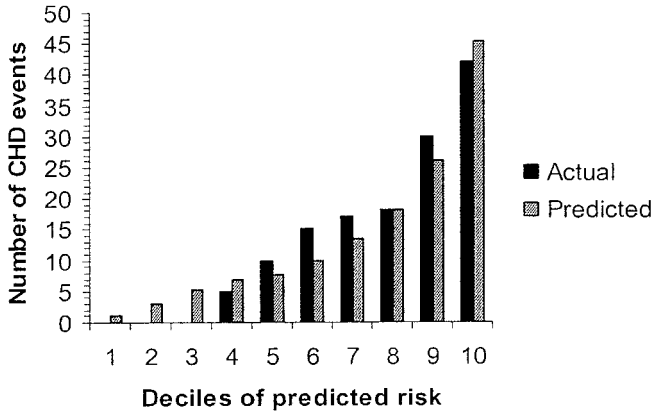

Figure 1—*Predicted and actual number of CHD events over 5 years in Salford by decile of predicted risk based on the Tayside model.*

CONCLUSIONS — This study provides the first population-derived model for prediction of risk of major CHD events in people with type 2 diabetes who are initially free of cardiovascular disease and validated on an independent dataset. The estimates are based on routine measurements of known risk factors for CHD, along with glycemic control and duration of diabetes, which are specific to type 2 diabetes. These factors were also found to be important in the UKPDS algorithm (19). The model also included height, which is consistent with the Prospective Cardiovascular Munster Study algorithm (12). The estimates are easily calculated from the Weibull accelerated failure time model compared with the Cox proportional hazards model (28). This model not only allows the estimation of survival probabilities and event rates for health service planning, and clinical trial design, but also allows the prediction of risk in individual patients to guide decision making on the use of cholesterol and blood pressure–lowering drugs.

The recent British Hypertension Society guidelines (29) for patients with type 2 diabetes and hypertension and American College of Physicians guidelines (30) for those with cardiovascular risk factors apply fairly blanket criteria for selection of statins, while an estimate of absolute risk of CHD allows a more focused guide to selection of lipid-lowering therapy. Before this study, most CHD predictions relied on the Framingham cohort containing only 337 people with diabetes (11–13), which only has a single indicator for the presence of diabetes. Several studies (1–3) have indicated a greater risk of CHD and long-term mortality in people with diabetes.

The Framingham equations also lack diabetes-specific factors such as A1C and duration of diabetes, which have been shown to be associated with risk of CHD (19). Validation studies (31–33) of the Framingham equations have yielded mixed results. Recently, a diabetes-specific equation was derived from a sample enrolled in the UKPDS randomized controlled trial, and they also found that A1C and duration of diabetes were important predictors of risk (19). However, the subjects in this study differed from the Tayside diabetic population cohort in a number of important ways. First, the mean age was 52 years compared with a mean age of 61 years in this study. Second, 60% of subjects in the UKPDS study were male compared with 52% male subjects in our population, the latter figure being more consistent with that found in diabetic population–based registers in the U.K. Third, baseline levels of A1C, blood pressure, and cholesterol were all higher in our population. Finally, perhaps the most important difference relates to the observed mortality. During the first 4 years of the UKPDS, the standardized mortality ratio was less than that observed in the general population at 0.94 and 0.96 for male and female subjects, respectively, possibly because patients with life-threatening illnesses were excluded from the study. This led to the UKPDS risk engine excluding individuals with <4 years follow-up as well as those aged >65 years, giving an equation conditional on surviving 4 years initially. In contrast, no one was excluded from the Tayside population cohort due to comorbidity, and all ages were included. Consequently, this algorithm can be applied to all patients with type 2 diabetes who are typical of patients seen in clinics throughout the U.K.

A1C is an important clinical measurement of glycemic control, which is generally available and which previously has been demonstrated an association of poor control with adverse outcomes (19). In addition, the interaction with time clearly demonstrates increased risk associated with poorer control, especially in the following 5 years, but it is clear in a population where A1C is being monitored that if a patient with poor glycemic control survives 5 years, baseline A1C is not a good predictor over the long term. However, from a patient and clinical perspective,

5-year predictions would be extremely useful. There is some evidence that patients with diabetes find difficulties in relating current behavior and medication with what are seen as possible events in the far future (34). Clearly, the communication of risk to influence behavior is also important (35).

Strength of the Tayside cohort is the relatively long follow-up period with a maximum of 9.5 years with good ascertainment of events. All prognostic equations may suffer from overfitting to the data from which they were derived and hence need to be validated on different populations (31). The algorithm had good discrimination; that is, the probability that the model gives a high risk to those who develop CHD in 5 years compared with those who do not was 0.71. For comparative purposes, the Framingham equation gave a discrimination of $c = 0.79$ (31). In addition, the algorithm derived in the Tayside population transported well to an independent population in Salford, England, with good calibration (36). A limitation of our study is that we did not include nonmajor coronary events such as angina. However, we concentrated on major coronary events for consistency with previous studies and for ease of transportability (19,37).

In conclusion, this study provides for the first time a type 2 diabetes–specific risk validated equation for CHD derived from an unselected population. It provides a useful tool for the clinician faced with the increasing prevalence of a chronic condition (38) associated with the high risk of CHD. It will aid decision making to provide early appropriate action to decrease the risk of adverse outcomes as well as aid health service planning and design of clinical trials.

Acknowledgments— This study was supported by a research grant from Diabetes UK and Pfizer UK.

We thank Douglas Boyle, who set up the DARTS database, and data facilitators Ritchie MacAlpine, Janice Broomhall, and Karen Hunter.

References

1. UK Prospective Diabetes Study Group: Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes (UKPDS 38). *BMJ* 317:703–713, 1998
2. Gray RP, Yudkin JS: Cardiovascular disease in diabetes mellitus. In *Textbook of Diabetes*. Pickup J, Williams G, Eds. Oxford, U.K., Blackwell Scientific, 1997, p. 57.1–57.22
3. Haffner SM, Lehto S, Ronnemaa T, Pyorala K, Laakso M: 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, 1998
4. Garcia MJ, McNamara PM, Gordon T, Kannel WB: Morbidity and mortality in diabetics in the Framingham population: sixteen year follow-up study. *Diabetes* 23: 105–111, 1974
5. Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, MacFarlane PW, McKillop JH, Packard CJ, the West of Scotland Coronary Prevention Study Group: Prevention of coronary heart disease with pravastatin in men with hypercholesterolaemia. *N Engl J Med* 333:1301–1307, 1995
6. Scandinavian Simvastatin Survival Study Group: Randomised trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 344:1383–1389, 1994
7. Heart Protection Study Collaborative Group: MRC/BHF Heart Protection Study of cholesterol-lowering with simvastatin in 5963 people with diabetes: a randomised placebo-controlled trial. *Lancet* 361: 2005–2016, 2003
8. Hansson L, Zanchetti A, Carruthers SG, Dahlof B, Elmfeldt D, Julius S, Menard J, Rahn KH, Wedel H, Westerling S: Effects of intensive blood-pressure lowering and low-dose asprin in patients with hypertension: principal results of the Hypertension Optimal Treatment (HOT) randomised trial. *Lancet* 351:1755–1762, 1998
9. Colwell JA: Aspirin therapy in diabetes. *Diabetes Care* 20:1767–1771, 1997
10. Jackson R, Lawes CM, Bennett DA, Milne RJ, Rodgers A: Treatment with drugs to lower blood pressure and blood cholesterol based on an individual's absolute cardiovascular risk. *Lancet* 365:434–441, 2005
11. Wood D, Durrington P, Poulter N, McInnes G, Rees A, Wray R: Joint British recommendations on prevention of coronary heart disease in clinical practice. *Heart* 80 (Suppl. 2):S1–S19, 1998
12. Schulte H, Assmann G: CHD risk equations obtained from the Framingham Heart Study applied to the PROCAM study of cardiovascular risk factors. *Lancet* 1:126–133, 1991
13. Anderson KM, Odell PM, Wilson PWF, Kannel WB: Cardiovascular disease risk profiles. *Am Heart J* 121:293–298, 1990
14. Voss R, Cullen P, Schulte H, Assmann G: Prediction of risk of coronary events in middle-aged men in the Prospective Cardiovascular Munster Study (PROCAM) using neural networks. *Int J Epidemiol* 31: 1253–1262, 2002
15. McEwan P, Williams JE, Griffiths JD, Bagnst A, Peters JR, Hopkinson P, Currie CJ: Evaluating the performance of the Framingham risk equations in a population with diabetes. *Diabet Med* 21:318–323, 2004
16. Isles C: Consensus statement: consensus conference on lipid lowering to prevent vascular events. *Proc Roy Coll Phys Edin* 29:1–3, 1999
17. Guideline Development Group: *Lipids and the Primary Prevention of Heart Disease*. Edinbutgh, U.K., Scottish Intercollegiate Guidelines Network (SIGN), 1999 (SIGN 40)
18. Guideline Development Group: *Secondary Prevention of Coronary Heart Disease Following Myocardial Infarction.* Edinburgh, U.K., Scottish Intercollegiate Guidelines Network (SIGN), 2000 (SIGN 41)
19. Stevens RJ, Kothari V, Adler AI, Stratton IM, Holman RR: The UKPDS risk engine: a model for the risk of coronary heart disease in type II diabetes (UKPDS 56). *Clin Sci* 101:671–679, 2001
20. Morris AD, Boyle DIR, MacAlpine R, Emslie-Smith A, Jung RT, Newton RW, MacDonald TM: The Diabetes Audit and Research in Tayside Scotland (DARTS) study: electronic record linkage to create a district diabetes register. *BMJ* 315:524–528, 1997
21. Evans JMM, McDevitt DG, MacDonald TM: The Tayside Medicines Monitoring Unit (MEMO): a record linkage study for pharmacovigilance. *Pharm Med* 9:177–184, 1995
22. New JP, Hollis S, Campbell F, McDowell D, Burns E, Dornan TL, Young RJ: Measuring clinical performance and outcomes from diabetes information systems: an observational study. *Diabetologia* 43:836–843,2000
23. New JP, Burns JA, Hollis S, Bennett J, Young RY: Improving ascertainment rates for macrovascular disease within electronic diabetes records using data linkage (Abstract). *Diabet Med* 19 (Suppl. 2):218, 2002
24. Anderson KM: A non-proportional hazards Weibull accelerated failure time model. *Biometrics* 47:281–288, 1991
25. Grambsch PM, Therneau TM: Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 81: 515–526, 1994
26. Pencina MJ, D'Agostino RB: Overall c as a measure of discrimination in survival analysis model specific population value and confidence interval estimation. *Statist Med* 23:2109–2123, 2004
27. May S, Hosmer DW: A simplified method of calculating an overall goodness-of-fit test for the Cox proportional hazards model. *Lifetime Data Anal* 4:109–120, 1998
28. Cox DR: Regression models and life-tables (with discussion). *J R Stat Soc Ser [Ser B]* 34:187–220, 1972

29. Williams B, Poulter NR, Brown MJ, Davis M, McInnes GT, Potter JF, Sever PS, Thom SMG, the BHS Guidelines Working Party, the British Hypertension Society: British Hypertension Society guidelines for hypertension management 2004 (BHS-IV): summary. *BMJ* 328:634–640, 2004
30. Snow V, Aronson MD, Hornbake ER, Mottur-Pilson C, Weiss KB, the Clinical Efficacy Assessment Subcommittee of the American College of Physicians: Lipid control in the management of type 2 diabetes mellitus: a clinical practice guideline from the American College of Physicians. *Ann Intern Med* 140:644–649, 2004
31. D'Agostino RB Sr, Grundy S, Sullivan LM, Wilson P, the CHD Risk Prediction Group: Validation of the Framingham Coronary Heart Disease Prediction Scores. *JAMA* 286:180–187, 2001
32. Diverse Populations Collaboration Group: Prediction of mortality from coronary heart disease among diverse populations: is there a common predictive function. *Heart* 88:222–228, 2002
33. Brindle P, Emberson J, Lampe F, Walker M, Whincup P, Fahey T, Ebrahim S: Predictive accuracy of the Framingham coronary heart disease in British men prospective cohort study. *BMJ* 327:1267, 2003
34. Wolpert HA, Anderson BJ: Management of diabetes: are doctors framing the benefits from the wrong perspective? *BMJ* 323:994–996, 2001
35. Edwards A, Elwyn G, Mulley A: Explaining risks: turning numerical data into meaningful pictures. *BMJ* 324:827–830, 2002
36. van Houwelingen HC: Validation, calibration, revision and combination of prognostic survival models. *Statist Med* 19:3401–3415, 2000
37. Liu J, Hong Y, D'Agostino RB, Wu Z, Wang W, Sun J, Wilson PWF, Kannel WB, Zhao D: Predictive value for the Chinese population of the Framingham CHD risk assessment tool compared with the Chinese Multi-Provincial Cohort Study. *JAMA* 291:2591–2599, 2004
38. King H, Aubert RE, Herman WH: Global burden of diabetes, 1995–2025: prevalence, numerical estimates, and projections. *Diabetes Care* 21:1414–1431, 1998