



Performance of Cardiovascular Disease Risk Scores in People Diagnosed With Type 2 Diabetes: External Validation Using Data From the National Scottish Diabetes Register

Diabetes Care 2018;41:2010–2018 | <https://doi.org/10.2337/dc18-0578>

Stephanie H. Read,¹ Merel van Diepen,² Helen M. Colhoun,³ Nynke Halbesma,¹ Robert S. Lindsay,⁴ John A. McKnight,⁵ David A. McAllister,⁶ Ewan R. Pearson,⁷ John R. Petrie,⁴ Sam Philip,⁸ Naveed Sattar,⁴ Mark Woodward,^{9,10,11} and Sarah H. Wild,¹ on behalf of the Scottish Diabetes Research Network Epidemiology Group

OBJECTIVE

To evaluate the performance of five cardiovascular disease (CVD) risk scores developed in diabetes populations and compare their performance to QRISK2.

RESEARCH DESIGN AND METHODS

A cohort of people diagnosed with type 2 diabetes between 2004 and 2016 was identified from the Scottish national diabetes register. CVD events were identified using linked hospital and death records. Five-year risk of CVD was estimated using each of QRISK2, ADVANCE (Action in Diabetes and Vascular disease: preterAx and diamicroN-MR Controlled Evaluation), Cardiovascular Health Study (CHS), New Zealand Diabetes Cohort Study (NZ DCS), Fremantle Diabetes Study, and Swedish National Diabetes Register (NDR) risk scores. Discrimination and calibration were assessed using the Harrell C statistic and calibration plots, respectively.

RESULTS

The external validation cohort consisted of 181,399 people with type 2 diabetes and no history of CVD. There were 14,081 incident CVD events within 5 years of follow-up. The 5-year observed risk of CVD was 9.7% (95% CI 9.6, 9.9). C statistics varied between 0.66 and 0.67 for all risk scores. QRISK2 overestimated risk, classifying 87% to be at high risk for developing CVD within 5 years; ADVANCE underestimated risk, and the Swedish NDR risk score calibrated well to observed risk.

CONCLUSIONS

None of the risk scores performed well among people with newly diagnosed type 2 diabetes. Using these risk scores to predict 5-year CVD risk in this population may not be appropriate.

Despite improvements through earlier diagnoses and improved treatments, cardiovascular disease (CVD) mortality and morbidity risk among people with type 2 diabetes remain markedly higher than in people without diabetes (1,2). The effect size depends on the subtype of CVD as well as age, sex, diabetes duration, ethnicity, and socioeconomic status (3–5).

Accurate CVD risk estimation in people with type 2 diabetes without established CVD can identify patients at high risk of developing CVD and can thus be used to

¹Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, Edinburgh, U.K.

²Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden, the Netherlands

³Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, U.K.

⁴Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, U.K.

⁵Metabolic Unit, Western General Hospital, Edinburgh, U.K.

⁶Institute of Health and Wellbeing, University of Glasgow, Glasgow, U.K.

⁷Division of Cardiovascular and Diabetes Medicine, University of Dundee, Dundee, U.K.

⁸Diabetes Research Unit, NHS Grampian, Aberdeen, U.K.

⁹The George Institute for Global Health, University of Oxford, Oxford, U.K.

¹⁰The George Institute for Global Health, University of New South Wales, New South Wales, Australia

¹¹Department of Epidemiology, Johns Hopkins University, Baltimore, MD

Corresponding author: Stephanie H. Read, stephanie.read@ed.ac.uk

Received 16 March 2018 and accepted 11 June 2018.

This article contains Supplementary Data online at <http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc18-0578/-/DC1>.

© 2018 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <http://www.diabetesjournals.org/content/license>.

guide appropriate treatment, for example, with statins, to illustrate to patients the likely effects of lifestyle choices and identify eligible participants for clinical trials. The U.K. clinical guideline network, the National Institute of Health and Clinical Excellence, recently updated its guidelines to advocate using the QRISK2 score (6), a risk score developed in the general population (7), to ascertain CVD risk in people with type 2 diabetes. Despite this recommendation, the performance of QRISK2 has not been independently externally validated in people with type 2 diabetes.

Several CVD risk scores have also been developed specifically for use among people with type 2 diabetes (8). While most of the earliest diabetes-specific CVD risk scores, such as the UK Prospective Diabetes Study (UKPDS) risk engine, have been extensively externally validated, many of the contemporary risk scores have not (8–10). Though one recent study did externally validate several contemporary risk scores (11), this study was limited by small sample sizes of the external validation cohorts, resulting in imprecise estimates of calibration and discrimination (12,13). In addition, few external validation studies have been conducted on statin-naïve participants.

Scotland maintains a national register of all patients with a diagnosis of type 2 diabetes, and this register can be linked to population-based hospitalization and mortality records. Consequently, this data source offers an opportunity to explore the performance of existing risk scores in a contemporary population of people with type 2 diabetes.

We evaluated the predictive performance of five diabetes-specific CVD risk scores in an external validation cohort of people with type 2 diabetes in Scotland and compared their performance with that of QRISK2.

RESEARCH DESIGN AND METHODS

Study Design and Participants

Data for these analyses were obtained from the population-wide Scottish Care Information–Diabetes (SCI-Diabetes) database. This dynamic clinical register was established in 2000 and is populated by patient data from primary care and hospital diabetes clinics. Outcome data were obtained from linkage to the Scottish Morbidity Record (SMR01), a national hospital admission data set, and death

registrations. Approval for generation and analysis of the linked data set was obtained from the Caldicott guardians of all Health Boards in Scotland, the Privacy Advisory Committee of the Information Services Division of NHS National Services Scotland, and the multicenter research ethics committee.

The external validation cohort consisted of people diagnosed with type 2 diabetes between 1 January 2004 and 1 June 2016 in Scotland. This time frame was chosen because SCI-Diabetes achieved >99% completeness of primary and secondary care clinics from 2004 onward. The cohort was restricted to people who had no previous history of CVD (as defined below) and who were aged between 30 and 89 years at date of diagnosis of diabetes owing to small numbers of people in other age-groups. We excluded people with a history of CVD at diagnosis of diabetes from our cohort since all but one of the risk scores that we wished to validate were designed to estimate risk of incident CVD. We included individuals who were prescribed statins prior to and after type 2 diabetes diagnosis in the main analyses but conducted sensitivity analyses in subpopulations restricted to 1) people who had not been prescribed statins prior to type 2 diabetes diagnosis and 2) people who had not been prescribed statins prior to type 2 diabetes diagnosis or during follow-up.

Members of the cohort were followed up from baseline, defined as date of diabetes diagnosis, until date of death, date of first CVD event, or study end date (1 June 2016)—whichever came first.

Outcome

CVD was defined as any hospital admission or death from myocardial infarction, stroke, unstable angina, transient ischemic attack, peripheral vascular disease, and coronary, carotid, or major amputation procedures between baseline and 1 June 2016. ICD-10 codes and Office of Population Censuses and Surveys Classification of Interventions and Procedures, version 4, codes used for identifying CVD may be found in the Supplementary Data.

Selected Risk Scores

QRISK2 was developed using data from the QResearch database, is based on a Cox proportional hazards model, and predicts 10-year risk of CVD (7). A previous systematic review identified 12

CVD risk prediction models designed for use among individuals with type 2 diabetes (9). Of these, five—the Swedish National Diabetes Register (NDR) risk score (14), the Action in Diabetes and Vascular disease: preterAx and diamicroN-MR Controlled Evaluation (ADVANCE) CVD risk score (15), the Fremantle Diabetes Study risk score (16), the New Zealand Diabetes Cohort Study (NZ DCS) risk score (17), and the Cardiovascular Health Study (CHS) risk score (18)—were chosen, as these were developed to predict CVD, while the remaining risk scores predict only coronary heart disease or stroke. Since the publication of the systematic review, an additional risk score for CVD, the Atherosclerosis Risk in Communities (ARIC) risk score, has been developed (19). However, this risk score includes several predictors (alcohol consumption and physical activity) that are not available in SCI-Diabetes and linked data sources and so was not considered in this validation exercise.

The characteristics of QRISK2 and the five diabetes-specific risk scores are presented in Table 1. All five diabetes-specific risk scores were derived from Cox proportional hazards models; three predict 5-year risk, while CHS predicts 10-year risk and ADVANCE predicts 4-year risk. The 5-year baseline hazard for QRISK2 has previously been published, while the 5-year baseline hazards were obtained from the study investigators for CHS and were estimated by extrapolation for ADVANCE.

Predictors Used in Risk Models

Taken together, the selected CVD risk prediction models contain the following predictors: age, sex, diabetes status (type 1/type 2/no diabetes), diabetes duration, ethnicity, Townsend deprivation score, systolic blood pressure, pulse pressure, smoking status, BMI, total-to-HDL cholesterol ratio, HDL cholesterol, non-HDL cholesterol, glycosylated hemoglobin, glucose-lowering medications, micro-/macroalbuminuria, albumin-to-creatinine ratio, creatinine, family history of CVD, antihypertensive medications, lipid-lowering medications, retinopathy, chronic kidney disease, rheumatoid arthritis, and atrial fibrillation.

Definitions of Predictors in External Validation Cohort

Baseline predictor values were defined as measurements recorded closest to baseline and no more than 24 months prior to

Table 1—Characteristics of QRISK2 and five CVD risk scores for use in people with type 2 diabetes

Name	Population	Cohort type	Time frame	Follow-up time, years	Main outcome	Risk factors	Internal validation C statistic (95% CI)
QRISK2 risk score (7)	2.3 million people, aged 35–74 years, in England and Wales without previous CVD	Electronic health records	1993–2008	Mean 7.3 for women, 6.9 for men	10-year risk of CHD, stroke, or transient ischemic attack (ICD-10 I20, I22–I25, I63–I64). Not peripheral arterial disease	Age, sex, diabetes status, ethnicity, BMI, total-to-HDL cholesterol ratio, SBP, atrial fibrillation, smoking, treated hypertension, Townsend social deprivation score, rheumatoid arthritis, family history of CHD	Men 0.792 (0.789, 0.794), women 0.817 (0.814, 0.820)
Swedish NDR risk score (14)	24,288 people, aged 30–74 years, in Sweden	Register	2002–2007	Mean 4.8	5-year risk of fatal or nonfatal CVD, nonfatal CHD (ICD-10, I20–I21), PCI or CABG, fatal CHD (I20–I25), or nonfatal or fatal stroke (I61, I63, I64)	Age, sex, diabetes duration, BMI, total-to-HDL cholesterol ratio, SBP, HbA _{1c} , smoking, treated hypertension, lipid-lowering drugs, micro- and macroalbuminuria, previous history of CVD	0.72
ADVANCE CVD risk score (15)	7,168 people, aged ≥55 years, without previous CVD from 215 collaborating centers in 20 countries from Asia, Australia, Europe, and North America	Trial	Recruitment 2001–2003	Mean 4.5	4-year risk of fatal or nonfatal MI or stroke or cardiovascular death. ICD-9 codes for nonfatal event 430–435, 437–438, 410; ICD-9 codes for fatal event 394–459, 798.9	Age, sex, diabetes duration, HbA _{1c} , atrial fibrillation, treated hypertension, albumin-to-creatinine ratio, pulse pressure, retinopathy, non-HDL cholesterol	0.70 (0.68, 0.73)
Fremantle Diabetes Study risk score (16)	1,240 people, mean age 64.1 years, from Fremantle, Western Australia	Observational cohort study	Recruitment 1993–1996, follow-up until 2006	Mean 4.5	5-year risk of fatal or nonfatal MI, stroke, or sudden death (no ICD codes provided)	Age, sex, ethnicity, prior CVD, HbA _{1c} , albumin-to-creatinine ratio, HDL cholesterol	0.80
NZ DCS risk score (17)	36,127 people, median age 59 years, without previous CVD from New Zealand	Observational cohort study	2000–2009	Median 3.9	First fatal or nonfatal CVD event and coronary and peripheral arterial procedures (see online Data Supplement in ref. 17)	Age, sex, diabetes duration, ethnicity, total-to-HDL cholesterol, SBP, HbA _{1c} , smoking, albuminuria	0.68
CHS risk score (18)	782 people, aged >65 years, without previous CVD from four field centers in the U.S.	Observational cohort study	Recruitment between 1989 and 1993, follow-up until 1999	Mean 7.0	10-year risk of MI, stroke, and death (no ICD codes provided)	Age, sex, smoking status, HbA _{1c} , SBP, total cholesterol, HDL cholesterol, creatinine, use of glucose-lowering medications	0.64

CABG, coronary artery bypass grafting; CHD, coronary heart disease; HbA_{1c}, glycated hemoglobin; MI, myocardial infarction; PCI, percutaneous coronary intervention; SBP, systolic blood pressure.

or 12 months after date of diagnosis of diabetes. Any predictor without a measurement within this time frame was declared missing. Prescriptions of antihypertensive and lipid-lowering medications occurring within the 3 months preceding baseline date were defined using British National Formulary codes 2.5 and 2.12, respectively. Chronic kidney disease was defined as a recording of estimated glomerular filtration rate of <60 mL/min/ 1.73 m² and/or a hospital admission for chronic kidney disease (ICD-10 codes N18 and I12–I13 and ICD-9 code 585).

Some predictors that appear within the risk scores were not available within SCI-Diabetes and linked data sets, and therefore some proxy predictors were used. Presence of rheumatoid arthritis was defined as patients with any prescription for disease-modifying antirheumatic drugs, defined with a British National Formulary code of 10.1.3 prior to baseline. Atrial fibrillation was defined as a hospital admission record, including diagnosis codes for atrial fibrillation (ICD-10 code I48 and ICD-9 code 427.3), or a warfarin prescription in the absence of a hospital record of prior deep vein thrombosis or pulmonary embolism (20). For area-based deprivation, the contemporary Scottish measure (Scottish Index of Multiple Deprivation [SIMD]) (21) was mapped across to the historical Townsend score (see Supplementary Table 1). Family history was estimated as the conditional probability of having a family history of CVD based on age and deprivation status (SIMD) with use of data from the 2014 Scottish Health Survey (see Supplementary Table 2) (22).

We conducted sensitivity analyses whereby all proxy categorical predictors (atrial fibrillation, rheumatoid arthritis, and family history of CVD) were set to null and where the Townsend score was set to the mean. Further sensitivity analyses were conducted to include prevalent diabetes whereby baseline was defined as the latest of 1 January 2010, date of diabetes diagnosis, or date of 30th birthday. Lastly, we examined whether the predictive performance of the selected risk scores changed over time (based on diabetes diagnosis before or during/after 2011).

Statistical Analyses

Missing predictor data were imputed using multiple imputation assuming data

were missing at random (*mice* package in R) (23). The imputation model included all predictors and the outcome (follow-up time and CVD event) and was used to generate 20 imputed data sets. Estimates were pooled using Marshall's adaptation of Rubin's rules (24). Complete case analyses were also conducted as additional sensitivity analyses.

Observed 5-year risk of CVD was estimated using the Kaplan-Meier estimator. Five-year risk of CVD was estimated at time of type 2 diabetes diagnosis using each of the five selected CVD risk scores and QRISK2. The predictive performance of the selected risk scores was assessed by examining measures of calibration and discrimination. Calibration describes how closely the predicted 5-year risk and the observed 5-year risk agree and was assessed by plotting smoothed observed incidence by predicted incidence using Kaplan-Meier estimates (25). Calibration-in-the-large statistics and calibration slopes for which values of 0 and 1, respectively, indicate good calibration were also calculated. Calibration-in-the-large statistics compare the mean predicted risk and mean observed risks. Calibration statistics were also calculated for the recalibrated risk scores after adjustment of the baseline hazard to that of the external validation cohort (26,27). Discrimination describes the model's ability to differentiate patients who developed CVD from those who did not and was assessed here by calculating the Harrell *C* statistic. This statistic describes the probability that, for any pair of individuals among whom one developed CVD and the other did not develop CVD, the predicted risk of the outcome is higher for the individual who did subsequently develop the disease (28). A *C* statistic of 1 denotes perfect discrimination, and a value of 0.5 denotes a prediction model that performs no better than a flip of a coin.

We calculated the number of people classified as high risk, based on the cut-off point in national clinical guidelines ($\geq 10\%$ estimated risk in QRISK2), or low risk ($< 10\%$ estimated risk in QRISK2) (6).

All statistical analyses were carried out in R, version 3.2.2. Calibration plots were generated using the *rms* package in R (29). The reporting of this external validation study is in accordance with the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines (30).

RESULTS

There were 218,607 individuals diagnosed with type 2 diabetes in Scotland between January 2004 and June 2016 (Table 2). Of these, 37,208 had a previous history of CVD and were excluded from the analyses, leaving 181,399 individuals to form the external validation cohort. Of the 26 predictors included in the risk models, 11 had missing values, and the average missingness was 18%. There were a total of 118,098 individuals with incomplete predictor data, including 33,210 individuals with a single incomplete predictor and a further 42,834 individuals with two incomplete variables only (Supplementary Table 3).

Overall, there were 14,081 incident CVD events during 673,740 person-years of follow-up, and the 5-year observed Kaplan-Meier risk of CVD was 9.7% (95% CI 9.6, 9.9). The median follow-up time was 5 years, and there were 91,549 individuals who were followed up for at least 5 years. There were 10,023 non-CVD deaths during follow-up.

Within the external validation cohort, 36,471 individuals had been prescribed statins prior to date of diabetes diagnosis. During follow-up, 71,585 individuals were prescribed statins and the median time until statin initiation was 141 days.

Calibration and Discrimination

Measures of calibration and discrimination are presented in Table 3, and calibration plots are presented in Fig. 1. Briefly, the agreement between observed and predicted risks (calibration in the large) was better with use of the Swedish NDR, CHS, and NZ DCS risk scores than with the QRISK2 and ADVANCE risk scores. Overall, QRISK2 overestimated risk while ADVANCE underestimated risk across all risk groups. *C* statistics for each of the models ranged between 0.663 (95% CI 0.658, 0.668) and 0.674 (0.669, 0.679) for the whole population. These values decreased after stratification by age, particularly in older age-groups. Supplementary Fig. 1 presents the distribution of predicted risks for each risk score.

Risk Classification

With a 10% threshold for high risk of developing CVD, QRISK2 classified 86.8% of the cohort as high risk, capturing 13,633 (96.8%) of the subsequent CVD events. In comparison, 3.2, 58.8, 25.8, 82.6, and 37.3% of the cohort were classified as high risk, capturing 8.4, 80.8, 59.2, 94.7, and 46% of the CVD events

Table 2—Baseline characteristics of individuals diagnosed with type 2 diabetes in Scotland between 2004 and 2016 by subsequent 5-year CVD outcome status over a median follow-up of 5 years

	CVD event	No CVD events
<i>N</i>	14,081	167,318
Age at diagnosis, years, median (IQR)	66.5 (17.4)	59.3 (18.0)
Sex, <i>n</i> (%)		
Men	8,292 (8.4)	90,604 (91.6)
Women	5,789 (7.0)	76,714 (93.0)
Ethnicity, <i>n</i> (%)		
White	9,808 (7.6)	118,633 (92.4)
Southeast Asian	220 (5.4)	3,836 (94.6)
Other	480 (6.0)	7,579 (94.0)
SIMD, <i>n</i> (%)		
Most deprived	3,700 (8.4)	40,349 (91.6)
2	3,361 (8.2)	37,867 (91.8)
3	2,780 (7.6)	33,569 (92.4)
4	2,463 (7.5)	30,576 (92.5)
Least deprived	1,777 (6.6)	24,957 (93.4)
Systolic blood pressure, mmHg, mean (SD)	139.9 (19.9)	138.6 (17.7)
Pulse pressure, mmHg, mean (SD)	60.4 (16.3)	57 (14.7)
Smoking status, <i>n</i> (%)		
Current smoker	3,854 (9.7)	35,946 (90.3)
Ex-smoker	5,463 (8.9)	56,232 (91.1)
Never smoker	4,699 (5.9)	74,493 (94.1)
BMI, kg/m ² , mean (SD)	31.3 (6.5)	32.9 (6.9)
Total-to-HDL cholesterol ratio, mean (SD)	4.5 (1.6)	4.7 (1.6)
Non-HDL cholesterol, mmol/mol, mean (SD)	3.9 (1.3)	4.1 (1.3)
Glycated hemoglobin, mmol/L, mean (SD)	64.0 (23.0)	64.8 (23.4)
Glycated hemoglobin, %, mean (SD)	8.0 (4.1)	8.1 (4.2)
Albuminuria, <i>n</i> (%)		
Normal	5,664 (6.6)	80,735 (93.4)
Microalbuminuria	1,466 (9.3)	14,333 (90.7)
Macroalbuminuria	215 (13.6)	1,361 (86.4)
Albumin-to-creatinine ratio, mean (SD)	5.3 (18.2)	3.3 (12.4)
Prescribed antihypertensive medications, <i>n</i> (%)		
Yes	6,053 (9.3)	58,958 (90.7)
No	8,028 (6.9)	108,360 (93.1)
Prescribed rheumatoid arthritis medications, <i>n</i> (%)		
Yes	210 (9.7)	1,946 (90.3)
No	13,871 (7.7)	165,372 (92.3)
Atrial Fibrillation, <i>n</i> (%)		
Yes	1,487 (17.3)	7,098 (82.7)
No	12,594 (7.3)	160,220 (92.7)
Retinopathy, <i>n</i> (%)		
Yes	1,735 (9.2)	17,068 (90.8)
No	12,346 (7.6)	150,250 (92.4)
Chronic kidney disease, <i>n</i> (%)		
Yes	3,547 (13.6)	22,454 (86.4)
No	9,712 (6.7)	135,537 (93.3)
Prescribed statins prior to diabetes diagnosis, <i>n</i> (%)		
Yes	4,509 (12.4)	31,962 (87.6)
No	9,572 (6.6)	135,356 (93.4)

using the ADVANCE, CHS, Fremantle, NZ DCS, and Swedish NDR risk scores, respectively (Supplementary Table 4).

Sensitivity Analyses

After recalibration of the risk scores, calibration improved slightly for the

ADVANCE risk score (Supplementary Fig. 2). The agreement between observed and predicted risks estimated by QRISK2 deteriorated further. The median predicted risk estimated by the recalibrated QRISK2, ADVANCE, CHS, Fremantle Diabetes Study, NZ DCS,

and Swedish NDR risk scores was 94.7, 2.5, 4.7, 4.1, 6.5, and 6.4%, respectively.

Among the subset of individuals who were not prescribed statins prior to diabetes diagnosis (*n* = 144,928), there were 9,572 events during 533,006 person-years of follow-up. Measures of calibration and discrimination for this subset yielded results similar to those of the main analyses (Supplementary Fig. 3 and Supplementary Table 5). These findings were also replicated in the subset of individuals who were not prescribed statins prior to diabetes diagnosis or during follow-up (Supplementary Fig. 4 and Supplementary Table 5), when proxy variables were replaced with null or mean values (Supplementary Fig. 5 and Supplementary Table 5), when people with prevalent diabetes were included in the cohort (Supplementary Fig. 6 and Supplementary Table 6), and when complete case analyses were used with omission of missing data (Supplementary Fig. 7 and Supplementary Table 5). The predictive performance of each of the risk scores varied only slightly depending on year of diabetes diagnosis (<2011 vs. ≥2011) (Supplementary Table 5).

CONCLUSIONS

Using a population-wide diabetes dataset, we have conducted the largest external validation of several contemporary CVD risk scores and the first external evaluation of QRISK2 among people with type 2 diabetes.

The ability of the assessed risk scores to discriminate between people who did and did not develop incident CVD as assessed by Harrell *C* statistics was similar, with all *C* statistics for all risk scores <0.68. The median predicted risk using QRISK2 was 23.5% compared with an observed risk of 9.3%, and QRISK2 classified >86% of people with type 2 diabetes as high risk. Compared with QRISK2, the agreement between predicted and observed risks with use of the risk scores developed in diabetes populations was generally better. For example, the median predicted risk with use of the CHS and Swedish NDR risk scores was 11.7 and 8.3%, respectively. The ADVANCE risk score exhibited the poorest calibration and severely underestimated risk of CVD in people with type 2 diabetes in Scotland. Recalibration by adjustment of the baseline hazard worsened the performance of QRISK2, since the 5-year baseline hazard

Table 3—Age-stratified calibration and discrimination statistics for QRISK2 and five diabetes-specific risk scores

Risk score	Age-group (years)	Observed 5-year risk	Predicted 5-year risk, %, median (IQR)	Calibration in the large	Calibration slope	C statistic (discrimination)
QRISK2	Overall	9.7	24.07 (21.21)	−0.14	0.376 (0.376, 0.377)	0.674 (0.669, 0.679)
	30–45	3.4	8.73 (9.71)	−0.06	0.208 (0.208, 0.208)	0.666 (0.644, 0.689)
	46–60	6.8	18.26 (13.81)	−0.11	0.272 (0.272, 0.273)	0.632 (0.623, 0.641)
	61–75	11.5	29.51 (16.54)	−0.19	0.317 (0.317, 0.317)	0.604 (0.597, 0.612)
	>75	21.0	45.01 (17.55)	−0.24	0.374 (0.374, 0.375)	0.578 (0.568, 0.588)
ADVANCE	Overall	9.7	2.00 (2.53)	0.08	1.808 (1.805, 1.811)	0.666 (0.661, 0.671)
	30–45	3.4	0.58 (0.45)	0.02	3.283 (3.277, 3.289)	0.628 (0.605, 0.651)
	46–60	6.8	1.33 (0.96)	0.06	2.353 (2.350, 2.356)	0.595 (0.586, 0.605)
	61–75	11.5	2.93 (2.09)	0.08	1.657 (1.655, 1.660)	0.594 (0.587, 0.602)
	>75	21.0	6.27 (4.57)	0.15	0.973 (0.970, 0.976)	0.575 (0.565, 0.585)
CHS	Overall	9.7	11.71 (11.17)	−0.02	0.631 (0.631, 0.632)	0.674 (0.669, 0.679)
	30–45	3.4	4.58 (2.92)	−0.02	0.760 (0.759, 0.760)	0.638 (0.615, 0.661)
	46–60	6.8	8.68 (5.34)	−0.02	0.742 (0.742, 0.742)	0.622 (0.613, 0.632)
	61–75	11.5	16.1 (9.64)	−0.05	0.546 (0.545, 0.547)	0.603 (0.596, 0.611)
	>75	21.0	26.17 (15.56)	−0.05	0.398 (0.396, 0.400)	0.575 (0.565, 0.585)
Fremantle	Overall	9.7	5.24 (7.63)	0.05	0.738 (0.737, 0.738)	0.665 (0.660, 0.670)
	30–45	3.4	1.2 (0.88)	0.02	2.025 (2.023, 2.027)	0.626 (0.603, 0.648)
	46–60	6.8	3.23 (2.2)	0.04	1.157 (1.156, 1.159)	0.591 (0.582, 0.600)
	61–75	11.5	8.49 (5.52)	0.03	0.736 (0.735, 0.736)	0.593 (0.585, 0.600)
	>75	21.0	20.63 (12.11)	0.00	0.497 (0.496, 0.497)	0.580 (0.570, 0.590)
NZ DCS	Overall	9.7	16.17 (10.87)	−0.06	0.725 (0.725, 0.725)	0.670 (0.665, 0.674)
	30–45	3.4	7.7 (2.9)	−0.05	0.679 (0.676, 0.683)	0.645 (0.622, 0.667)
	46–60	6.8	12.67 (4.37)	−0.06	0.740 (0.739, 0.741)	0.609 (0.599, 0.618)
	61–75	11.5	20.23 (6.39)	−0.09	0.725 (0.725, 0.726)	0.599 (0.591, 0.606)
	>75	21.0	30.45 (8.42)	−0.09	0.635 (0.633, 0.638)	0.573 (0.563, 0.583)
Swedish NDR	Overall	9.7	8.26 (6.79)	0.02	0.955 (0.954, 0.955)	0.663 (0.658, 0.668)
	30–45	3.4	3.67 (2.26)	−0.01	0.871 (0.871, 0.871)	0.632 (0.609, 0.654)
	46–60	6.8	6.44 (3.56)	0.01	0.869 (0.869, 0.870)	0.602 (0.592, 0.611)
	61–75	11.5	10.54 (5.62)	0.00	0.727 (0.727, 0.727)	0.589 (0.582, 0.596)
	>75	21.0	16.79 (8.74)	0.04	0.576 (0.575, 0.576)	0.566 (0.556, 0.575)

of the external validation study was higher than the 5-year baseline hazard in the QResearch development data set. More advanced recalibration approaches, in which regression coefficients of the predictors are adjusted, are required to ensure better agreement between QRISK2-predicted and observed risks in people with type 2 diabetes in Scotland (6). The poor performance of QRISK2 among people with type 2 diabetes could lead to the overtreatment of people at low risk.

Findings from Other Studies

Although U.K. national clinical guidelines recommend the use of QRISK2 to estimate CVD risk in people with type 2 diabetes, the performance of QRISK2 in estimating CVD risk in external populations has not previously been assessed. However, an evaluation of the performance of QRISK2 in people with type 2 diabetes has been made using a subset of people with type 2 diabetes in the QResearch database and is described

in an online report (31). This approach to validation, whereby the performance of the model was assessed in a subset of the derivation cohort, is likely to have led to optimistic measures of performance. As expected, therefore, the C statistics describing the discriminative ability of QRISK2 were better in this evaluation than in our validation (C statistics 0.703 [95% CI 0.691, 0.715] in women and 0.696 [0.685, 0.706] in men), while the agreement between predicted and observed risks was also better.

Most previous studies have reported that CVD risk scores developed in general populations underestimate risk in people with type 2 diabetes (8), so we were surprised to find that QRISK2 overestimated risk in our external validation cohort. However, this difference may be partly explained by the inclusion of patients with prevalent type 2 diabetes in the QRISK2 derivation cohort, though sensitivity analyses in which people with prevalent type 2 diabetes were included

in the external validation cohort did not markedly improve the performance of QRISK2 (Supplementary Fig. 6). Inclusion of diabetes in the risk score as a categorical variable and in an interaction with age as in this risk score and others is unlikely to sufficiently capture the complex relationship between diabetes and CVD, particularly the effect of diabetes duration on CVD risk. Similarly, prediction of CVD risk in people with type 2 diabetes is likely to be further complicated by the possible presence of type 2 diabetes subtypes with distinct disease trajectories (32). Identification of whether the incorporation of variables denoting type 2 diabetes subtypes within existing risk scores would improve their performance would be of interest for future research.

Previous validation studies of contemporary diabetes-specific risk scores are limited. One recent external validation study assessed the performance of the five diabetes-specific risk scores in three

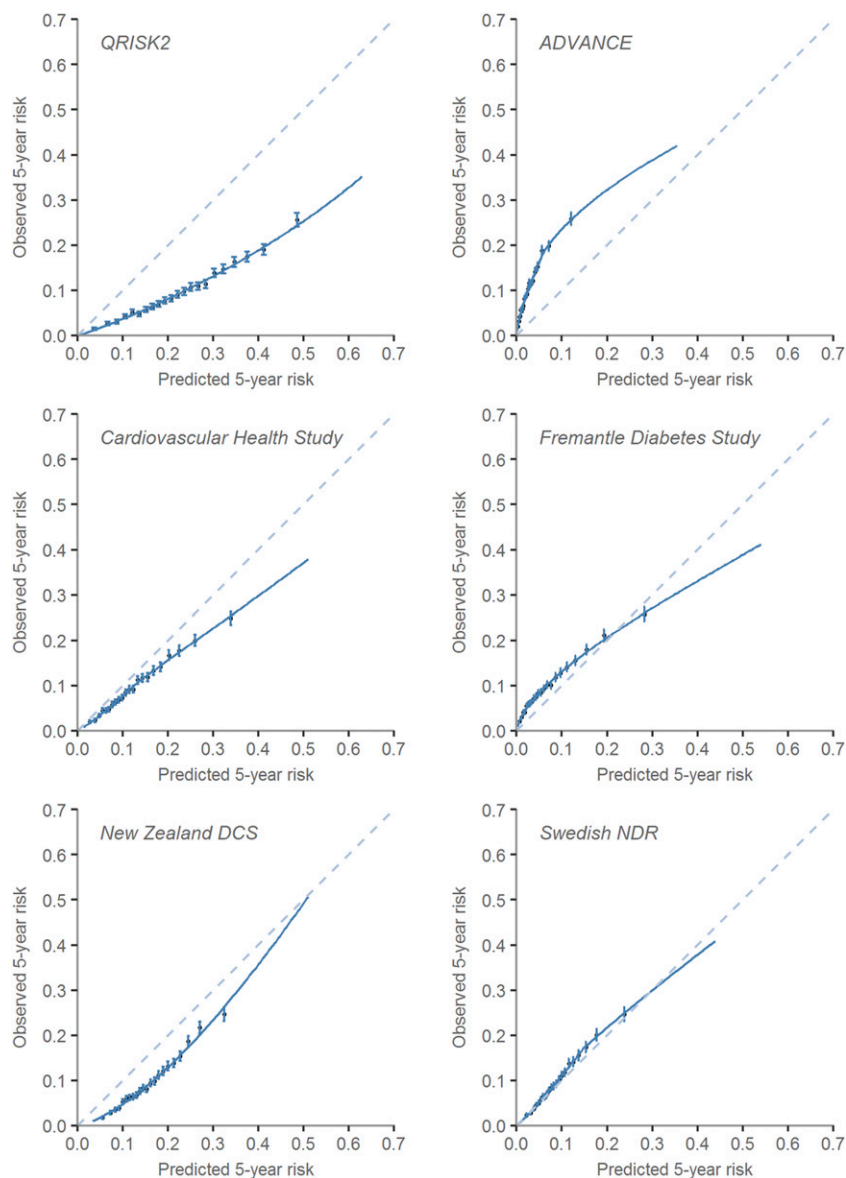


Figure 1—Calibration plots for observed versus predicted 5-year risk of CVD as estimated using QRISK2, ADVANCE, CHS, Fremantle Diabetes Study, NZ DCS, and Swedish NDR risk scores in people diagnosed with type 2 diabetes between 2004 and 2016 in Scotland. Gray dashed line reflects perfect agreement between observed and predicted risk.

separate cohorts; the European Prospective Investigation into Cancer and Nutrition (EPIC-NL, EPIC-Potsdam, and the Secondary Manifestations of ARterial disease (SMART) study (11). Expected risk (according to each respective risk score) to observed risk ratios varied between 1.06 (95% CI 0.81, 1.40) and 1.46 (1.04, 2.05). The risk scores exhibited poor discriminative ability in the three external validation cohorts with *C* statistics ranging from 0.54 (95% CI 0.46, 0.63) for the CHS risk score in EPIC-NL to 0.69 (0.59, 0.79) for the Fremantle risk score in SMART. Within

each external validation cohort, the discriminative ability was similar for each risk score, a finding replicated in the current study and a possible reflection of the limitations of the Harrell *C* statistic in the presence of extensive censoring (28). Unfortunately, the wide CI owing to the small numbers of events in each external validation cohort (52 events in EPIC-NL, 73 in EPIC-Potsdam, and 58 in SMART) made interpretation of the performance of these models difficult and prevented the authors from identifying the strongest performing risk score. The ADVANCE risk score was externally validated in

1,836 patients enrolled in the DIABHYCAR (noninsulin-dependent diabetes, hypertension, microalbuminuria or proteinuria, cardiovascular events, and ramipril) clinical trial and exhibited discrimination similar (*C* statistic 0.69 [95% CI 0.65, 0.72]) to that reported here, but it underestimated risk in the DIABHYCAR population (15).

Beyond differences in the performance of different health systems, there are likely to be a number of explanations for the overall poor performance of the assessed risk scores (33). One major potential explanation is differences in the distribution of outcomes and predictors (i.e., the case mix) in the external validation cohort compared with the derivation cohorts. Different age distributions are likely to be the most important difference between development cohorts and this external validation cohort, as indicated by the age-stratified measures of discrimination and calibration in Table 3. A further factor that may have contributed to poor performance in this cohort is different eligibility criteria. For example, ADVANCE was a trial with strict inclusion criteria that made for a very nonstandard population (34). Definitions of CVD also varied between derivation and validation cohorts. While QRISK2 identifies angina through general practice records, the current study only includes hospital admissions for angina and therefore angina incidence will be underestimated. Other factors that may have contributed to the poor performance of these risk scores were the use of proxies, different time frames of the outcome (10-year development vs. 5-year validation), and, potentially, differences in patterns of glucose-lowering therapies that may have different effects on CVD risks.

Strengths/Weaknesses

This study had a number of strengths. By utilizing population-based registers we were able to assemble the largest external validation cohort of people with type 2 diabetes to assess and directly compare the performance of several CVD risk scores to date. The large cohort also enabled the assessment of each model's performance in subsets of people based on statin exposure. The population-based nature of these data also ensured low risk of selection biases influencing our findings and enabled us to present results

that are applicable to the entire population of Scotland.

A number of weaknesses of the study should be acknowledged. Firstly, the use of proxy measures for some of the predictor variables may have contributed to the poor performance of the models for which these were required. However, by conducting sensitivity analyses to explore the likely effect of using these proxy measures, we have shown that this limitation is unlikely to have had a large effect on the overall findings of our study. Concerns surrounding the accuracy of the recording of CVD events may be a further limitation of this work. Nonetheless, findings from the West of Scotland Coronary Prevention Study (WOSCOPS) indicated that linkage to hospital admissions registers for acquiring CVD events may be as effective as direct patient contact (35). Finally, we were unable to validate all existing risk scores for people with type 2 diabetes owing to the unavailability of some predictors, though risk scores that include variables that are generally not measured may be difficult to implement in clinical practice. We acknowledge that further research is needed to establish whether diabetes treatment contributes to CVD risk independently of other factors. Such research will be particularly valuable for new diabetes drugs that appear to have a beneficial effect on CVD in trial populations.

Implications/Conclusions

Risk scores have important roles in guiding treatment, communicating risks to patients, and identifying eligible clinical trial participants. Unfortunately, we have shown that many existing risk scores do not accurately predict incident CVD risk in people with newly diagnosed type 2 diabetes, though risk scores developed in diabetes populations generally performed better than QRISK2. Current guidelines that recommend using QRISK2 would classify 87% of people with type 2 diabetes in Scotland as high risk, leading to the potential overtreatment of individuals at low risk. This approach is therefore not dissimilar to classifying all people aged over 40 years and with type 2 diabetes as high risk, as recommended in several existing clinical guidelines (36–38).

We conclude that there is scope to improve risk scores for incident CVD among people with type 2 diabetes and suggest that QRISK2 and the five

diabetes-specific risk scores, without recalibration, do not currently meet the standard for application to real-world patients in Scotland.

Funding. Funding for this project came from the Chief Scientist Office (PDF/15/07). D.A.M. is funded via an Intermediate Clinical Fellowship and Beit Fellowship from the Wellcome Trust (201492/Z/16/Z).

The funding source had no role in the design, execution, analysis, or interpretation of this study.

Duality of Interest. H.M.C. reports grants, personal fees, and nonfinancial support from AstraZeneca LP, Boehringer Ingelheim, Bayer, Eli Lilly, Novartis Pharmaceuticals, Regeneron, Pfizer, Roche Pharmaceuticals, sanofi-aventis, Sanofi, and Novo Nordisk outside the submitted work. R.S.L. reports personal fees from Novo Nordisk, Eli Lilly, and Servier outside the submitted work. E.R.P. reports personal fees from Eli Lilly, Merck Sharp & Dohme, and Novo Nordisk outside the submitted work. S.P. reports personal fees from AstraZeneca, Sanofi, and Napp Pharmaceuticals outside the submitted work. N.S. reports grants and personal fees from Boehringer Ingelheim, Janssen, Novo Nordisk, Eli Lilly, Amgen, AstraZeneca, and Sanofi outside the submitted work. M.W. reports personal fees from Amgen outside the submitted work. No other potential conflicts of interest relevant to this article were reported.

Author Contributions. S.H.R. carried out data preparation and statistical analyses and wrote the first draft of the paper. The study was conceived by S.H.R., M.v.D., N.H., D.A.M., M.W., and S.H.W. S.H.R., M.v.D., H.M.C., N.H., R.S.L., J.A.M., D.A.M., E.R.P., J.R.P., S.P., N.S., M.W., and S.H.W. contributed to the interpretation of the findings and the manuscript's critical revision. S.H.R., M.v.D., H.M.C., N.H., R.S.L., J.A.M., D.A.M., E.R.P., J.R.P., S.P., N.S., M.W., and S.H.W. approved the final version of the manuscript. S.H.R. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Prior Presentation. Parts of this study were presented in abstract form at Diabetes UK Professional Conference 2018, London, U.K., 14–16 March 2018, and at the Annual Meeting of the European Diabetes Epidemiology Group 2018, Copenhagen, Denmark, 21–24 April 2018.

References

- Hansen MB, Jensen ML, Carstensen B. Causes of death among diabetic patients in Denmark. *Diabetologia* 2012;55:294–302
- Sarwar N, Gao P, Seshasai SR, et al.; Emerging Risk Factors Collaboration. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies [published correction appears in *Lancet* 2010;376:958]. *Lancet* 2010; 375:2215–2222
- Jackson CA, Jones NRV, Walker JJ, et al.; Scottish Diabetes Research Network Epidemiology Group. Area-based socioeconomic status, type 2 diabetes and cardiovascular mortality in Scotland. *Diabetologia* 2012;55:2938–2945

- Shah AD, Langenberg C, Rapsomaniki E, et al. Type 2 diabetes and incidence of cardiovascular diseases: a cohort study in 1.9 million people. *Lancet Diabetes Endocrinol* 2015;3:105–113
- Woodward M, Peters SA, Huxley RR. Diabetes and the female disadvantage. *Womens Health (Lond)* 2015;11:833–839
- National Institute for Health and Care Excellence (NICE). Cardiovascular disease: risk assessment and reduction, including lipid modification [Internet], clinical guideline [CG181], 2014. Available from <https://www.nice.org.uk/Guidance/cg181>. Accessed 15 March 2018
- Hippisley-Cox J, Coupland C, Vinogradova Y, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;336:1475–1482
- Chamnan P, Simmons RK, Sharp SJ, Griffin SJ, Wareham NJ. Cardiovascular risk assessment scores for people with diabetes: a systematic review. *Diabetologia* 2009;52:2001–2014
- van Dieren S, Beulens JW, Kengne AP, et al. Prediction models for the risk of cardiovascular disease in patients with type 2 diabetes: a systematic review. *Heart* 2012;98:360–369
- van Dieren S, Peelen LM, Nöthlings U, et al. External validation of the UK Prospective Diabetes Study (UKPDS) risk engine in patients with type 2 diabetes. *Diabetologia* 2011;54:264–270
- van der Leeuw J, van Dieren S, Beulens JW, et al. The validation of cardiovascular risk scores for patients with type 2 diabetes mellitus. *Heart* 2015;101:222–229
- Collins GS, Ogundimu EO, Altman DG. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Stat Med* 2016;35:214–226
- Vergouwe Y, Steyerberg EW, Eijkemans MJ, Habbema JD. Substantial effective sample sizes were required for external validation studies of predictive logistic regression models. *J Clin Epidemiol* 2005;58:475–483
- Zethelius B, Eliasson B, Eeg-Olofsson K, Svensson AM, Gudbjörnsdóttir S, Cederholm J; NDR. A new model for 5-year risk of cardiovascular disease in type 2 diabetes, from the Swedish National Diabetes Register (NDR). *Diabetes Res Clin Pract* 2011;93:276–284
- Kengne AP, Patel A, Marre M, et al.; ADVANCE Collaborative Group. Contemporary model for cardiovascular risk prediction in people with type 2 diabetes. *Eur J Cardiovasc Prev Rehabil* 2011;18:393–398
- Davis WA, Knuiam MW, Davis TM. An Australian cardiovascular risk equation for type 2 diabetes: the Fremantle Diabetes Study. *Intern Med J* 2010;40:286–292
- Elley CR, Robinson E, Kenealy T, Bramley D, Drury PL. Derivation and validation of a new cardiovascular risk score for people with type 2 diabetes: the New Zealand Diabetes Cohort Study. *Diabetes Care* 2010;33:1347–1352
- Mukamal KJ, Kizer JR, Djoussé L, et al. Prediction and classification of cardiovascular disease risk in older adults with diabetes. *Diabetologia* 2013;56:275–283
- Parrinello CM, Matsushita K, Woodward M, Wagenknecht LE, Coresh J, Selvin E. Risk prediction of major complications in individuals with diabetes: the Atherosclerosis Risk in Communities Study. *Diabetes Obes Metab* 2016;18:899–906

20. Morley KI, Wallace J, Denaxas SC, et al. Defining disease phenotypes using national linked electronic health records: a case study of atrial fibrillation. *PLoS One* 2014;9:e110900
21. Scottish Government. *Scottish Index of Multiple Deprivation 2012: A National Statistics Publication for Scotland*. Edinburgh, U.K., Scottish Government, 2012
22. Brown L, Christie S, Gill V, et al. *The Scottish Health Survey 2014: A National Statistics Publication for Scotland*. Edinburgh, U.K., Scottish Government, 2014
23. van Buuren S, Groothuis-Oudshoorn K. Mice: multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1–67
24. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009;9:57
25. Steyerberg EW. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*. New York, NY, Springer, 2009
26. Crowson CS, Atkinson EJ, Therneau TM. Assessing calibration of prognostic risk scores. *Stat Methods Med Res* 2016;25:1692–1706
27. Moons KG, Kengne AP, Grobbee DE, et al. Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012;98:691–698
28. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. *Stat Med* 2011;30:1105–1117
29. Harrell FEJ. rms: Regression Modeling Strategies. R package version 5.1–2 [Internet], 2018. Available from <https://CRAN.R-project.org/package=rms>. Accessed 15 March 2018
30. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement [published correction appears in *Ann Intern Med* 2015;162:600]. *Ann Intern Med* 2015;162:55–63
31. Hippisley-Cox J, Coupland C, Brindle P. Validation of QRISK2 (2014) in patients with diabetes [article online], 2014. Available from <http://eprints.nottingham.ac.uk/3602/1/QRISK2-2014-validation-in-patients-with-diabetes-1.7-web.pdf>. Accessed 15 March 2018
32. Ahlqvist E, Storm P, Käräjämäki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *Lancet Diabetes Endocrinol* 2018;6:361–369
33. Riley RD, Ensor J, Snell KI, et al. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ* 2016;353:i3140
34. Patel A, MacMahon S, Chalmers J, et al.; ADVANCE Collaborative Group. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *N Engl J Med* 2008;358:2560–2572
35. The West of Scotland Coronary Prevention Study Group. Computerised record linkage: compared with traditional patient follow-up methods in clinical trials and illustrated in a prospective epidemiological study. *J Clin Epidemiol* 1995;48:1441–1452
36. Piepoli MF, Hoes AW, Agewall S, et al.; ESC Scientific Document Group. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts). Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). *Eur Heart J* 2016;37:2315–2381
37. American Diabetes Association. 8. Cardiovascular disease and risk management. *Diabetes Care* 2015;38(Suppl.):S49–S57
38. Scottish Intercollegiate Guidelines Network (SIGN). Risk estimation and the prevention of cardiovascular disease: a national clinical guideline [Internet], 2017. SIGN publication no. 149, Edinburgh. Available from <http://www.sign.ac.uk/sign-149-risk-estimation-and-the-prevention-of-cardiovascular-disease.html>. Accessed 15 March 2018