ESC European Society of Cardiology

European Heart Journal (2023) **44**, 2544–2556
https://doi.org/10.1093/eurheartj/ehad260

CLINICAL RESEARCH
Epidemiology, prevention, and health care policies

SCORE2-Diabetes: 10-year cardiovascular risk estimation in type 2 diabetes in Europe

SCORE2-Diabetes Working Group and the ESC Cardiovascular Risk Collaboration*†

Received 10 June 2022; revised 6 April 2023; accepted 17 April 2023; online publish-ahead-of-print 29 May 2023

See the editorial comment for this article 'Risk prediction in patients with diabetes: is SCORE 2D the perfect solution?', by L. Rydén *et al.*, https://doi.org/10.1093/eurheartj/ehad263.

Abstract

Aims To develop and validate a recalibrated prediction model (SCORE2-Diabetes) to estimate the 10-year risk of cardiovascular disease (CVD) in individuals with type 2 diabetes in Europe.

Methods and results SCORE2-Diabetes was developed by extending SCORE2 algorithms using individual-participant data from four large-scale datasets comprising 229 460 participants (43 706 CVD events) with type 2 diabetes and without previous CVD. Sex-specific competing risk-adjusted models were used including conventional risk factors (i.e. age, smoking, systolic blood pressure, total, and HDL-cholesterol), as well as diabetes-related variables (i.e. age at diabetes diagnosis, glycated haemoglobin [HbA1c] and creatinine-based estimated glomerular filtration rate [eGFR]). Models were recalibrated to CVD incidence in four European risk regions. External validation included 217 036 further individuals (38 602 CVD events), and showed good discrimination, and improvement over SCORE2 (C-index change from 0.009 to 0.031). Regional calibration was satisfactory. SCORE2-Diabetes risk predictions varied several-fold, depending on individuals' levels of diabetes-related factors. For example, in the moderate-risk region, the estimated 10-year CVD risk was 11% for a 60-year-old man, non-smoker, with type 2 diabetes, average conventional risk factors, HbA1c of 50 mmol/mol, eGFR of 90 mL/min/1.73 m^2, and age at diabetes diagnosis of 60 years. By contrast, the estimated risk was 17% in a similar man, with HbA1c of 70 mmol/mol, eGFR of 60 mL/min/1.73 m^2, and age at diabetes diagnosis of 50 years. For a woman with the same characteristics, the risk was 8% and 13%, respectively.

Conclusion SCORE2-Diabetes, a new algorithm developed, calibrated, and validated to predict 10-year risk of CVD in individuals with type 2 diabetes, enhances identification of individuals at higher risk of developing CVD across Europe.

* Authors are listed at the end of the paper.
† Investigators of contributing studies are listed at the end of the paper.

Structured Graphical Abstract

Key Question

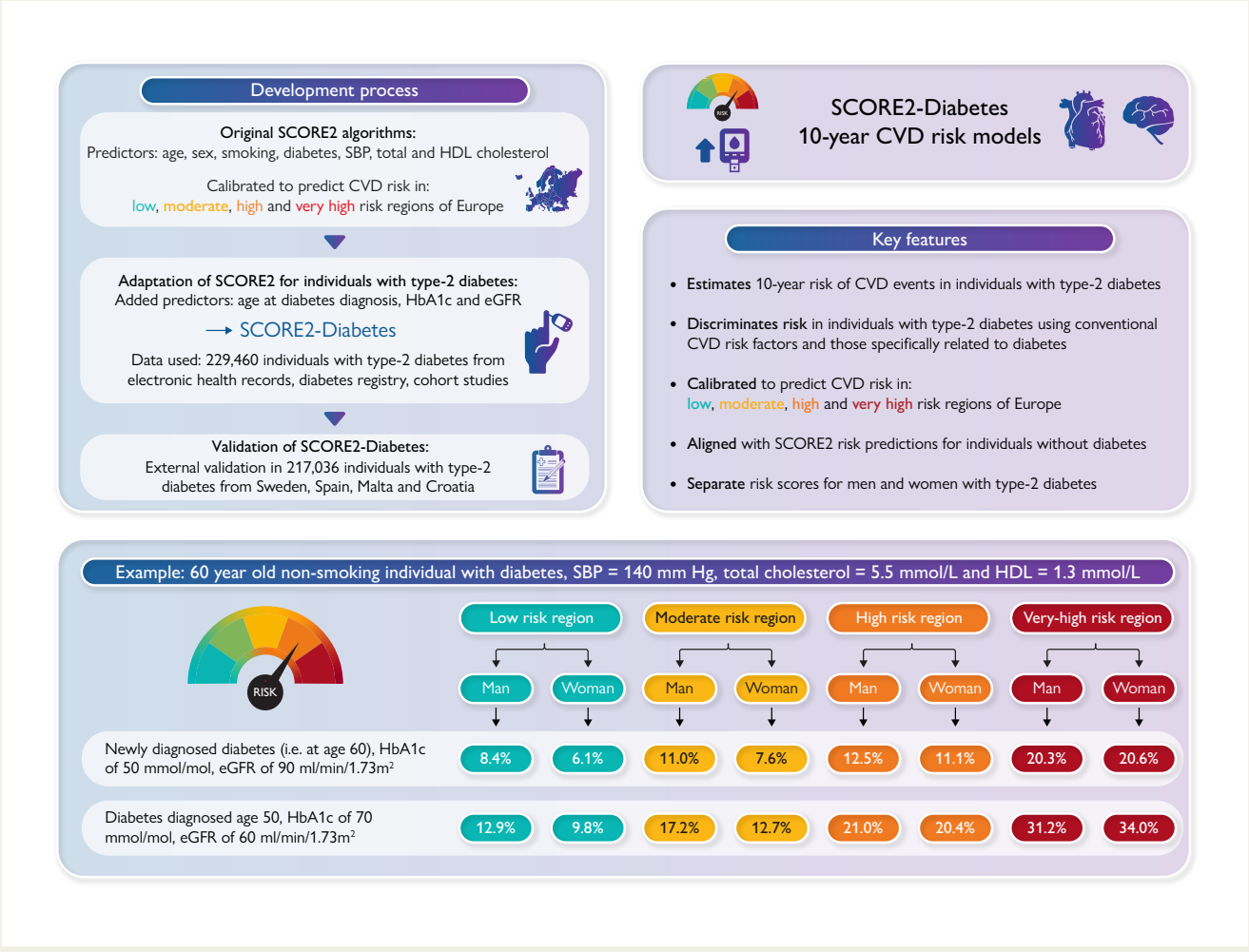
Can cardiovascular disease (CVD) risk prediction for individuals with type-2 diabetes be improved to reflect substantial regional variation in CVD incidence across Europe?

Key Finding

The SCORE2-Diabetes algorithms were developed by extending SCORE2, using data from >220,000 individuals with type 2 diabetes. Recalibration accounted for three- to four-fold variation in CVD incidence across Europe. SCORE2-Diabetes showed good external validation in >210,000 individuals from four countries (Sweden, Spain, Malta and Croatia).

Take Home Message

SCORE2-Diabetes accurately estimates CVD risk in individuals with type-2 diabetes. It extends SCORE2, aligning CVD risk prediction for those with and without diabetes, while accounting for variation in risk across Europe. This facilitates the identification of individuals at high CVD risk.



SCORE2-Diabetes 10-year CVD risk models: development process, key features and illustrative example. CVD: cardiovascular disease; SBP: systolic blood pressure; HDL-C: high-density lipoprotein cholesterol; HbA1c (mmol/mol): glycated haemoglobin, in International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) units; eGFR: estimated Glomerular Filtration Rate (mL/min/1.73m^2).

Keywords Prediction model • Diabetes • Cardiovascular diseases

Introduction

Cardiovascular diseases (CVD) remain a major cause of morbidity and mortality in Europe with almost 13 million new cases recorded in 2019 alone.[1] Type 2 diabetes mellitus is a major risk factor for CVD. Individuals with diabetes from high-income countries have, on average, a 2-fold greater risk of developing CVD outcomes compared to counterparts without diabetes.[2] The European Society of Cardiology (ESC) provides guidelines and advocates estimation of CVD risk in individuals with type 2 diabetes to inform treatment decisions.[3]

Risk prediction models used in the primary prevention of CVD in general populations usually estimate individual risk over a 10-year period by integrating information on measured levels of conventional CVD risk factors (i.e. age, smoking status, systolic blood pressure, and total and HDL-cholesterol) and information on diabetes status.[4–6] To help account for substantial variation in risk across individuals with diabetes, however, additional diabetes-related information [e.g. age at diagnosis of diabetes, glycated haemoglobin (HbA1c), and markers of kidney function] have been included in several published risk models.[7–10] Nonetheless, available diabetes-specific models have important potential limitations. In particular, they may not be optimal for use across Europe's diverse populations since they have been developed from a narrow set of observational studies and/or intervention trials, and have not been systematically 'recalibrated' (i.e. statistically adapted) to reflect the substantial variation in CVD rates across different European countries.[1,10,11] To address these limitations, the ESC has convened an effort to extend the regionally recalibrated European SCORE2 10-year risk models,[12] enabling use in individuals with type 2 diabetes.

Here, we describe development, validation, and illustration of SCORE2-Diabetes to estimate the 10-year risk of non-fatal myocardial infarction, stroke, or any CVD mortality in individuals with diabetes but without previous CVD, aged over 40 years, in four different European risk regions.

Methods

Study design

The SCORE2-Diabetes project involved several interrelated components and data sources (*Figure 1*). First, the original SCORE2 risk prediction models for fatal and non-fatal CVD outcomes were adapted for use in individuals with type 2 diabetes using individual-participant data from four population data sources [Scottish Care Information—Diabetes (SCID), Clinical Practice Research Datalink (CPRD), UK Biobank (UKB), Emerging Risk Factors Collaboration (ERFC)] across seven countries (England, Wales, Scotland, France, Germany, Italy, and the USA). Second, we recalibrated the derived risk models to each European risk region, applying methods previously used to develop SCORE2. Third, we completed external validation in individuals with type 2 diabetes across four countries (Sweden, Spain, Croatia, and Malta) using data from the Swedish National Diabetes Register (SNDR), the Information System for Research in Primary Care (SIDIAP, Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària), and two contributing registries from the EUropean Best Information through Regional Outcome in Diabetes (EUBIROD). Fourth, we illustrated the variation of CVD risk in individuals with type 2 diabetes across European regions by applying the recalibrated models to data from contemporary populations in each risk region.

Data sources and procedures

For model derivation, we used individual-participant data from patients with type 2 diabetes, without previous CVD, aged over 40 years, from SCID, CPRD, UKB, and seven cohorts from the ERFC with available information

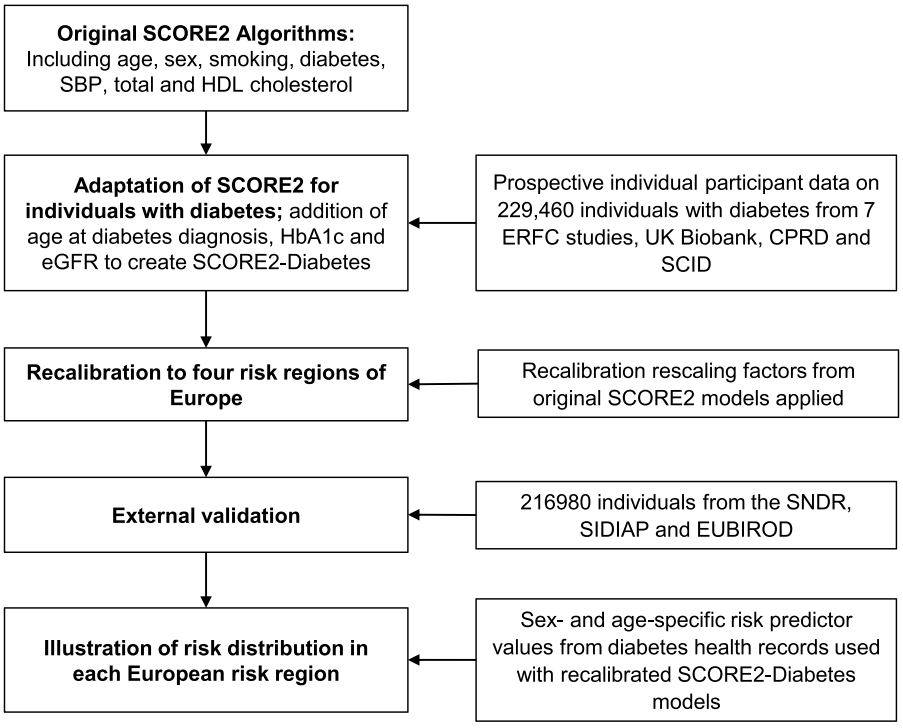


Figure 1 Study design for the SCORE2-Diabetes project. ERFC: Emerging Risk Factors Collaboration, CPRD: Clinical Practice Research Datalink, SCID: Scottish Care Information—Diabetes, SNDR: Swedish National Diabetes Register, SIDIAP: Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, EUBIROD: EUropean Best Information through Regional Outcome in Diabetes, eGFR: estimated Glomerular Filtration Rate, HbA1c: glycated haemoglobin.

on diabetes-related variables. SCID is a dynamic population-based register of people with a diagnosis of diabetes in Scotland that has had almost complete coverage since 2006.[13] CPRD is an ongoing primary care database of anonymised medical records from general practitioners, with coverage of over 11.3 million patients from 674 practices in the UK.[14] With 4.4 million active (alive, currently registered) patients meeting quality criteria, approximately 6.9% of the UK population are included and patients are broadly representative of the UK general population in terms of age, sex, and ethnicity. The data used for this study is restricted to the region of England. Model derivation datasets for the SCID and the CPRD involved individuals with diabetes on 1st June 2008 and risk factor measurements recorded during the period from 30th June 2006 to 31st December 2008. Follow-up was to 1st June 2019 for SCID and 31st December 2019 for CPRD, with incident non-fatal events obtained from linkage with Scottish Morbidity Records and English Hospital Episode Statistics and deaths from National Records of Scotland and Office for National Statistics. The UKB is a single large prospective cohort study with individual-participant data on approximately 500 000 participants aged over 40 years recruited across 23 UK-based assessment centres during 2006–10, and followed-up for cause-specific morbidity and mortality through linkages to routinely available national datasets and disease-specific registers.[15] The ERFC has collated and harmonised individual-participant data from many long-term prospective cohort studies of CVD risk factors and outcomes.[16] Prospective studies in the ERFC were included in this analysis if they met all the following criteria: had recorded baseline information on CVD risk factors necessary to derive risk prediction models [i.e. age, sex, smoking status, systolic blood pressure, total and HDL-cholesterol, history of diabetes mellitus (defined by self-report plus medication and/or biochemical criteria,[2,17]) age at diabetes diagnosis, HbA1c and creatinine or estimated glomerular filtration rate (eGFR)]; were approximately population-based [i.e. did not select participants based on having previous disease (e.g. case-control studies) and were not active treatment arms of intervention studies]; had a median year of baseline survey after 1990; and had recorded cause-specific deaths and/or non-fatal CVD events (i.e. non-fatal myocardial infarction or stroke) for at least 5 years of follow-up. Data selection for model adaptation/derivation is shown in Supplementary data online, *Figure S1*. Details of contributing data sources are provided in Supplementary data online, *Tables 1 and 2*.

For the recalibration of models, recalibration factors from the SCORE2 risk models were used. SCORE2 has been systematically recalibrated to reflect risk of the entire population (including those with diabetes) in four risk regions of Europe. Hence, adapting SCORE2 for use in individuals with type 2 diabetes (i.e. SCORE2-Diabetes) does not require additional data and recalibration for diabetes-specific populations. Data from the SNDR, SIDIAP, and EUBIROD were used for external validation (see Supplementary data online, *Table S3*). SNDR is a national registry that has close to complete coverage of the population with a diagnosis of type 2 diabetes in Sweden.[18] As with data used in model derivation, we used records from individuals with diabetes during the period from 30th June 2006 to 31st December 2008, and no previous history of CVD. Follow-up was to 31st December 2019 with incident fatal and non-fatal events obtained from linkage to hospital and mortality records. SIDIAP is a primary care electronic health records database managed by the Catalan Health Institute, covering around 75% of individuals (>5 million) in the Catalonia region of Spain across 328 primary care centres, and is representative of this population in terms of age, sex, and geographic distribution.[19,20] For this analysis, we used individuals with type 2 diabetes from a randomly selected 400 000 individuals whose records were linked to hospital and specific cause of death records to obtain CVD outcomes. Individuals had been included in SIDIAP for at least 1 year before 1st January 2010 and were subsequently followed-up until 2017. EUBIROD is the largest network of diabetes registries and data sources in Europe,[21] sharing a common dataset[22] and open source software[23] to analyse individual data in a privacy-enhanced distributed infrastructure.[24–26] Data on people with type 2 diabetes with baseline records between January 2013 and June 2015 were independently processed at each of the eight participating countries (Belgium, Croatia, Denmark, Germany, Hungary, Latvia, Malta, and Slovenia), and analysed using R source code embedded in the EUBIROD NeuBIRO software. Where available, follow-up for CVD events was obtained through linkage to hospital and death records over the subsequent 5 years, enabling validation. Only aggregate data were made available by each participating centre to the study coordinators. Risk factor data from CPRD, SNDR, SIDIAP, EUBIROD, and the 2017/18 extraction from the National Diabetes Audit (NDA) were used to illustrate SCORE2-Diabetes predicted risk distributions in each European risk region. The NDA is an annually updated registry covering more than 98% of individuals with a recorded diabetes diagnosis from primary healthcare providers in England and Wales and specialist healthcare providers in England.[27]

The primary outcome was CVD events, defined as a composite of cardiovascular mortality, non-fatal myocardial infarction, and non-fatal stroke. Follow-up was until the first non-fatal myocardial infarction, non-fatal stroke, death or end of the study, or registration period. Deaths from non-CVD were treated as competing events. Details of the different ICD-10 codes included in both the fatal and non-fatal components of the endpoint are provided in Supplementary data online, *Table S4*. In all data sources, individuals with a known history of previous CVD at baseline were excluded, as defined in Supplementary data online, *Table S5*.

Statistical analysis

Details of statistical analysis are provided in Supplementary data online, Supplementary Methods. For model derivation, the SCORE2 models were extended by addition of diabetes-related variables: HbA1c, age at diabetes diagnosis, and eGFR. These predictors were selected due to their predictive ability based on previous literature as well as their wide availability in clinical practice and available datasets used for model derivation. Coefficients for the variables already included in SCORE2 derivation (i.e. age, current smoking, history of diabetes mellitus, systolic blood pressure, and total and HDL-cholesterol) were fixed at the same values used in the SCORE2 models and included as an offset in Fine and Gray competing risk-adjusted models used to estimate additional sex-specific coefficients [i.e. sub-distribution hazard ratios (SHRs)]. Additional coefficients were then estimated for each of the SCORE2 variables, to allow their effects to vary among individuals with diabetes, as well as for the newly added diabetes-related variables included in SCORE2-Diabetes. All newly derived coefficients were estimated separately by data source and pooled using fixed effects meta-analysis. Since previous research showed that associations of these variables with CVD decline with increasing age, age interactions were added for all predictors. A quadratic term was also included for eGFR to allow for its non-linear association with CVD outcomes (see Supplementary data online, *Supplementary Methods Figure*). There were no (or very minimal) violations of the proportional hazards assumptions, as assessed by inclusion of time varying coefficients.

Risk models were recalibrated to risk regions using recalibration factors previously derived for SCORE2 and SCORE2-OP models (see Supplementary data online, *Supplementary Methods Table S1*). Similarly, the grouping of European countries into risk regions was defined according to World Health Organization CVD mortality rates following SCORE2 and SCORE2-OP methodology (see Supplementary data online, *Table S6 and Figure S2*). For validation, we assessed discrimination using Harrell's C-index, adjusted for competing risk,[28] and examined improvement when comparing use of SCORE2-Diabetes vs. SCORE2. Where data were available we compared SCORE2-Diabetes with the ADVANCE risk model for individuals with diabetes.[10] We use ADVANCE as a comparison as it is recommended by the 2021 ESC Guidelines on CVD prevention in clinical practice[3] and it is designed to predict CVD risk. To provide clinical context, we compared incremental improvements afforded by diabetes-related information included in SCORE2-Diabetes with those afforded by total and HDL-cholesterol, biomarkers commonly used in CVD risk assessment. Improvements in risk prediction were also quantified by the continuous net reclassification index (NRI), which summarises the appropriate directional change in risk predictions for those who do (cases) and do not (non-cases) experience an event during follow-up (with increases in

Table 1 Summary of available data on individuals with diabetes used in SCORE2-Diabetes risk model derivation

	N (%) or mean (SD)		
	SCID	ERFC/UKB	CPRD
Participants	136 192	20 517	72 751
Male sex	72 525 (53%)	11 485 (56%)	38 599 (53%)
SCORE2 variables			
Age (years)	65 (11)	60 (8)	64 (11)
Current smoker	24 447 (18%)	2353 (12%)	11 423 (21%)
Systolic blood pressure (mmHg)	136 (16)	142 (17)	136 (16)
Total cholesterol (mmol/L)	4.3 (1.0)	4.7 (1.1)	4.4 (1.0)
HDL-cholesterol (mmol/L)	1.3 (0.4)	1.2 (0.3)	1.2 (0.4)
SCORE2-Diabetes additional variables			
Diabetes age of diagnosis (per 5-years)	58 (12)	53 (9)	58 (11)
HbA1c (mmol/mol)	58 (17)	55 (20)	52 (19)
eGFR (mL/min/1.73 m^2)	74 (20)	88 (17)	76 (17)
Follow-up [years, median (5th–95th percentile)]	10.9 (6.8, 11.0)	11.3 (2.8, 13.6)	6.0 (0.8, 11.0)
Cardiovascular events	34 595	1864	7247
Non-cardiovascular deaths	21 062	1953	5211

SCID, Scottish Care Information—Diabetes; ERFC, Emerging Risk Factors Collaboration; UKB, UK Biobank; CPRD, Clinical Practice Research Datalink; eGFR, estimated Glomerular Filtration Rate, calculated using the CKD-EPI 2009 equations; HbA1c, glycated haemoglobin, in International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) units. Table shows summary statistics for datasets before imputation (which was carried out during analysis). A summary of missing data, by data source and variable, is provided in Supplementary data online, *Table S2*.

predicted risk being appropriate for cases and decreases being appropriate for non-cases). Similarly, the categorical NRI was also applied to summarise the appropriate movement between risk categories of <5%, 5%–10%, 10%–15%, 15%–20%, and >25%. Calibration was assessed by comparing the observed and predicted risks.

To compare the proportion of the population with diabetes at different levels of CVD risk according to the SCORE2-Diabetes models, predicted risk distributions were estimated using age- and sex-specific risk factor values from the CPRD, NDA, SNDR, and all contributing EUBIROD populations, with the region-specific recalibrated versions of SCORE2-Diabetes. To ensure that the SCORE2 recalibration factors were applicable in recalibration of SCORE2-Diabetes we assessed that the average sex- and age-specific SCORE2-Diabetes risk predictions matched the expected risks for each risk region, and that the average sex- and age-specific risk predictions were similar in the whole population, as well as in individuals with diabetes when using SCORE2 and SCORE2-Diabetes. In studies with available information, SHRs and observed absolute risks were also estimated using an extended endpoint additionally including non-fatal heart failure (HF) and peripheral artery disease (PAD) (see Supplementary data online, *Table S4*). We also ensured similar risk predictions were obtained when using both the 2009 and 2021[29] versions of the Chronic Kidney Disease Epidemiology Collaboration eGFR equations to ensure interchangeability of the two measures in clinical practice. Finally, SHRs were additionally estimated without the inclusion of ERFC/UKB data to ensure no sensitivity to potential minor overlap in individuals contributing to UK-based studies and the CPRD.

Missing data were imputed for derivation datasets, SNDR, and SIDIAP using methods described in the Supplementary data online, *Supplementary Methods*. We adopted analytical approaches and reporting standards recommended by the PROBAST guidelines and TRIPOD.[30] Analyses were performed with R-statistic programming (version 4.0.3, R Foundation for Statistical Computing, Vienna, Austria) and Stata (version 16.1, StataCorp, College Station, Texas). The study was designed and completed by the SCORE2-Diabetes Working Group in collaboration with the ESC Cardiovascular Risk Collaboration. Data used for the current study are available upon reasonable request and approval of the individual cohorts or collaborative groups. Stata code for calculation of the SCORE2-Diabetes algorithms is available on request from authors.

Results

Model derivation involved a total of 229 460 participants with diabetes and without history of CVD at baseline from SCID, CPRD, and ERFC/UKB. Mean age (SD) at baseline was 65 (11) years for SCID, 64 (11) years for CPRD, and 60 (8) years for ERFC/UKB. A total of 122 609 (53.4%) participants were male across all data sources (*Table 1*). Median (5th, 95th percentile) follow-up in years was 10.9 (6.8, 11.0) in SCID, 6.0 (0.8, 11.0) in CPRD, and 11.3 (2.8, 13.6) in ERFC/UKB, during which a total of 43 706 CVD events and 28 226 non-CVD deaths were recorded. SHRs are shown in *Table 2*. The association of the diabetes-related variables decreased with increasing age of participants (see Supplementary data online, *Supplementary Methods Figure*). Associations were similar when excluding ERFC/UKB data (see Supplementary data online, *Table S7*), and when an extended CVD endpoint including non-fatal HF and PAD was used (see Supplementary data online, *Table S8*).

The C-indices in the derivation datasets were 0.704 (95% CI 0.701, 0.706), 0.733 (0.727–0.739), and 0.666 (0.653, 0.678) in SCID, CPRD, and ERFC/UKB, respectively (*Figure 2*). In external validation, the C-index for SCORE2-Diabetes was 0.670 (0.667, 0.673) using data

Table 2 Sub-distribution hazard ratios for predictor variables in the SCORE2-Diabetes risk models

	Men		Women	
	Main effect	Age interaction term	Main effect	Age interaction term
SCORE2 variables				
Age (per 5 years)	1.71 (1.66, 1.76)	–	1.94 (1.88, 2.00)	–
Current smoking	1.61 (1.53, 1.70)	0.94 (0.91, 0.96)	1.85 (1.73, 1.98)	0.89 (0.87, 0.92)
Systolic blood pressure (per 20 mmHg)	1.14 (1.11, 1.17)	0.97 (0.96, 0.99)	1.15 (1.12, 1.19)	0.98 (0.97, 1.00)
Total cholesterol (per 1 mmol/L)	1.12 (1.10, 1.14)	0.98 (0.97, 0.99)	1.12 (1.09, 1.15)	0.98 (0.97, 0.99)
HDL-cholesterol (per 0.5 mmol/L)	0.90 (0.86, 0.93)	1.01 (0.99, 1.03)	0.85 (0.82, 0.89)	1.02 (1.00, 1.04)
History of diabetes mellitus	1.91 (1.81, 2.01)	0.91 (0.88, 0.93)	2.25 (2.11, 2.40)	0.88 (0.85, 0.91)
SCORE2-Diabetes additional variables				
Diabetes age at diagnosis (per 5-years)	0.90 (0.89, 0.91)		0.89 (0.88, 0.90)	–
HbA1c (per SD mmol/mol)	1.10 (1.09, 1.11)	0.99 (0.98, 0.99)	1.12 (1.11, 1.14)	0.98 (0.98, 0.98)
ln eGFR (per SD ln(mL/min/1.73m^2))	0.94 (0.93, 0.96)	1.01 (1.01, 1.01)	0.94 (0.92, 0.95)	1.02 (1.01, 1.02)
ln eGFR^2 (quadratic term)	1.01 (1.00, 1.01)	–	1.01 (1.00, 1.01)	–

Sex-specific sub-distribution hazard ratios from Fine and Gray models predicting the risk of fatal and non-fatal CVD events as derived for SCORE2 and adapted in individuals with diabetes from ERFC, UK Biobank, CPRD, SCID to include adjustments to SCORE2 effects and SCORE2-Diabetes additional variables. Age was centred at 60 years, systolic blood pressure at 120 mmHg, total cholesterol at 6 mmol/L, HDL-cholesterol at 1.3 mmol/L, age at diabetes onset at 50 years HbA1c at 31 mmol/mol and eGFR 90 mL/min/1.732 (i.e. ln-eGFR of 4.5). The median baseline survival at 10 years in the derivation cohorts was 0.9625 for men and 0.9795 for women. For HbA1c, 1 SD = 9.34 mmol/mol and for eGRF 1SD = 0.15 ln(mL/min/1.73 m^2).

Values shown are the combination of original SCORE2 coefficients and additional coefficients which modify the associations for individuals with diabetes. See Supplementary data online, *Supplementary methods* for full sets of component effects for each risk predictor.

from 168 585 individuals with diabetes (34 944 CVD events) from the SNDR and 0.658 (0.648, 0.669) using data from 21 698 individuals with diabetes (2464 CVD events) from SIDIAP. Using EUBIROD datasets including 3876 individuals from Malta and 22 821 individuals from Croatia with complete information on all risk predictors, the C-index was 0.661 (0.622, 0.699) and 0.688 (0.672, 0.705), respectively (see Supplementary data online, *Figure S3*).

In comparison to SCORE2, SCORE2-Diabetes showed improved risk discrimination in individuals with diabetes, with increases in C-indices (95% CI) of 0.021 (0.020, 0.022), 0.023 (0.020, 0.026), and 0.026 (0.018, 0.034) in SCID, CPRD, and ERFC/UKB, respectively. Somewhat smaller improvements were observed in SNDR, and SIDIAP with increases in C-index of 0.009 (0.007, 0.010) and 0.009 (0.005, 0.014), respectively (*Figure 2*). In EUBIROD datasets from Malta and Croatia, increases in C-indices were 0.031 (0.011, 0.050) and 0.013 (0.006, 0.021), respectively (see Supplementary data online, *Figure S3*). Significant improvements in C-indices were also seen in both men and women, and within 10-year age groups (see Supplementary data online, *Figures S4–S7*). C-indices were similar when eGFR was calculated using different equations (see Supplementary data online, *Figure S8*), but were slightly attenuated when excluding individuals with eGFR <45 mL/min/1.73 m^2 (see Supplementary data online, *Figure S9*). Improvements in risk discrimination provided by the additional diabetes-related variables included in SCORE2-Diabetes (i.e. age of diabetes diagnosis, HbA1c, and eGFR) were greater than that provided by total and HDL-cholesterol concentration in the same model. SCORE2-Diabetes also showed slightly improved discrimination over the ADVANCE risk score (see Supplementary data online, *Table S9*).

Using SCORE2-Diabetes rather than SCORE2 improved risk classification, yielding a continuous NRI of 25.2 (95% CI, 22.4, 28.0) in the CPRD and 28.7 (27.7, 29.8) in the SNDR. Similarly, using SCORE2-Diabetes rather than SCORE2 yielded a categorical NRI of 24.6 (22.5, 26.8) in the CPRD and 13.7 (12.9, 14.5) in the SNDR, with a respective net of 44.8% (43.0%, 46.7%) and 31.9% (31.2%, 32.6%) cases being appropriately reclassified (see Supplementary data online, *Table S10*).

After recalibration, the SCORE2-Diabetes predicted risks showed good agreement with the expected 10-year CVD incidence in each risk region (see Supplementary data online, *Figure S10*), and were similar on average within each age-group to those produced using SCORE2 (see Supplementary data online, *Figure S11*). SCORE2-Diabetes predicted risks also agreed with observed risks in individuals with diabetes from nationally representative datasets with 10-year of follow-up (see *Supplementary* data online, *Figure S12 and S13*), and showed improved calibration over SCORE2 (see Supplementary data online, *Figure S13*). Use of an extended CVD endpoint including non-fatal HF and PAD led to an absolute 10-year risk about 1.15 times higher than that estimated using the SCORE2-Diabetes CVD endpoint, with results varying slightly according to age (see Supplementary data online, *Figure S14*).

The SCORE2-Diabetes algorithms for CVD risk estimation in four European risk regions are shown in the Supplementary data online, *Supplementary Methods Table 1*. Risk charts to illustrate individual-specific estimation of 10-year CVD risk for are provided in Supplementary data online, *Appendix 1*, along with a risk calculator to give more precise individual estimates in Supplementary data online, *Appendix 2*. The estimated absolute risk for a given age and combination

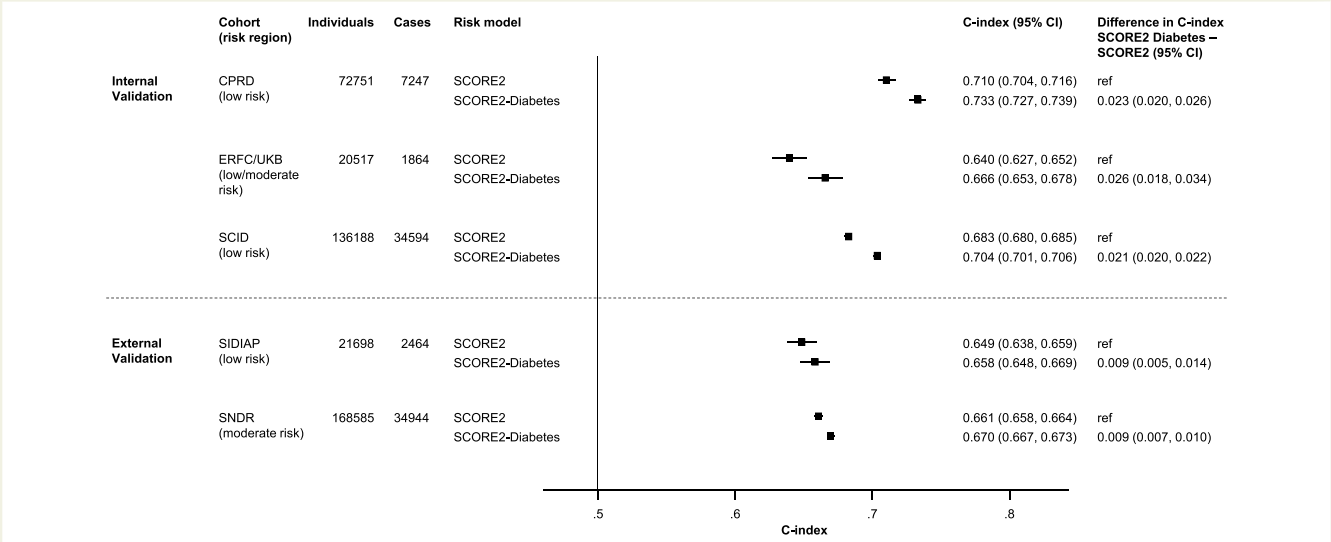


Figure 2 Internal and External validation of the SCORE2-Diabetes models: ability to discriminate CVD risk. CPRD: Clinical Practice Research Datalink; ERFC: Emerging Risk Factors Collaboration; UKB: UK Biobank; SCID: Scottish Care Information—Diabetes; SIDIAP: Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària; SNDR: Swedish National Diabetes Register.

of conventional CVD risk factors differed substantially according to levels of the diabetes-related variables (*Figure 3*). For example, when using the moderate-risk region version of SCORE2-Diabetes, the estimated 10-year CVD risk for a 60-year-old non-smoking man with a history of diabetes, average levels of conventional risk factors (i.e. systolic blood pressure of 140 mmHg, total cholesterol of 5.5 mmol/L, and HDL-cholesterol of 1.3 mmol/L), HbA1c of 50 mmol/mol, eGFR of 90 mL/min/1.73 m^2, and age at diabetes diagnosis of 60 years, was 11.0%. For a similar man with less favourable diabetes-related risk factors (i.e. HbA1c of 70 mmol/mol, eGFR of 60 mL/min/1.73 m^2, and age at diagnosis of 50 years), the estimated risks were 17.2%. For a woman with the same characteristics, risk was 7.9% and 12.7%, respectively. Risk estimates also varied across European risk regions due to recalibration, with a man or woman with the latter risk factor values having an estimated risk of 12.9% and 9.8%, respectively, in the low-risk region, and 31.2% and 34.0% in the very high-risk region (*Structured Graphical Abstract*, *Figure 3*).

When we applied recalibrated SCORE2-Diabetes models to simulated data representing populations from each risk region, the proportion of individuals aged 40–79 years with an estimated risk greater than 10% varied substantially by region, from 61% in the low-risk region to 96% in the very-high-risk region in men and from 51% to 94%, respectively, in women, with proportions increasing with age as expected (*Figure 4*).

Discussion

Compared with existing risk scores, SCORE2-Diabetes, an extension of the SCORE2 risk models tailored to individuals with type 2 diabetes across Europe's diverse populations, should help better support allocation of preventative interventions, as it involves several advantages.

First, SCORE2-Diabetes has been systematically recalibrated to four distinct European regions defined by varying CVD risk levels, using the most contemporary and representative CVD rates available.[12] This improves on previous CVD risk prediction models for individuals with diabetes which either have not incorporated any recalibration to different populations, or have been recalibrated based on sparse cohort or country-level data on individuals with diabetes, which may not accurately reflect the CVD rates and risk factor levels of populations in each region.[9–11] Our analysis illustrates that three- to four-fold variation in estimated CVD risk for a given set of risk factors can be seen as a result of recalibration. Without recalibration this substantial variation in risk across Europe would be ignored. Because the recalibration approach we used is based on registry data, the model can be readily updated to reflect future CVD incidence and risk factor profiles of any target population to be screened, including those with diabetes. This means that if descriptive age- and sex-specific epidemiological data are available from individual European countries (or within-country regions), they can be readily incorporated to revise models at a country-level. This is an important feature of the current risk score since there have been considerable changes in cardiovascular risk over time and region in people with type 2 diabetes, necessitating contemporary risk estimation.

A second – and related advantage – is that, rather than being developed solely in data from individuals with diabetes, SCORE2-Diabetes extends SCORE2 models that were developed in all individuals without previous CVD, including both those with and without diabetes (although the ESC does not recommend SCORE2 for use in those with diabetes). A key advantage of this approach is that it allows recalibration of the models using risk factor data and incidence rates from the general population, rather than requiring data specifically from individuals with diabetes, which are currently not available systematically across European countries. By extending SCORE2 we also ensure harmonisation of risk estimation for individuals with and without diabetes across Europe, aiding communication and interpretation of risk estimates. The existing ESC CVD Risk Calculation App[31] and the 'HeartScore' website[32] will be updated to include SCORE2-Diabetes to facilitate risk estimation and communication between health professionals and individuals with type 2 diabetes. Supplementary data online, *Appendices 1 and 2* also provide immediate tools to calculate individual risk estimates.

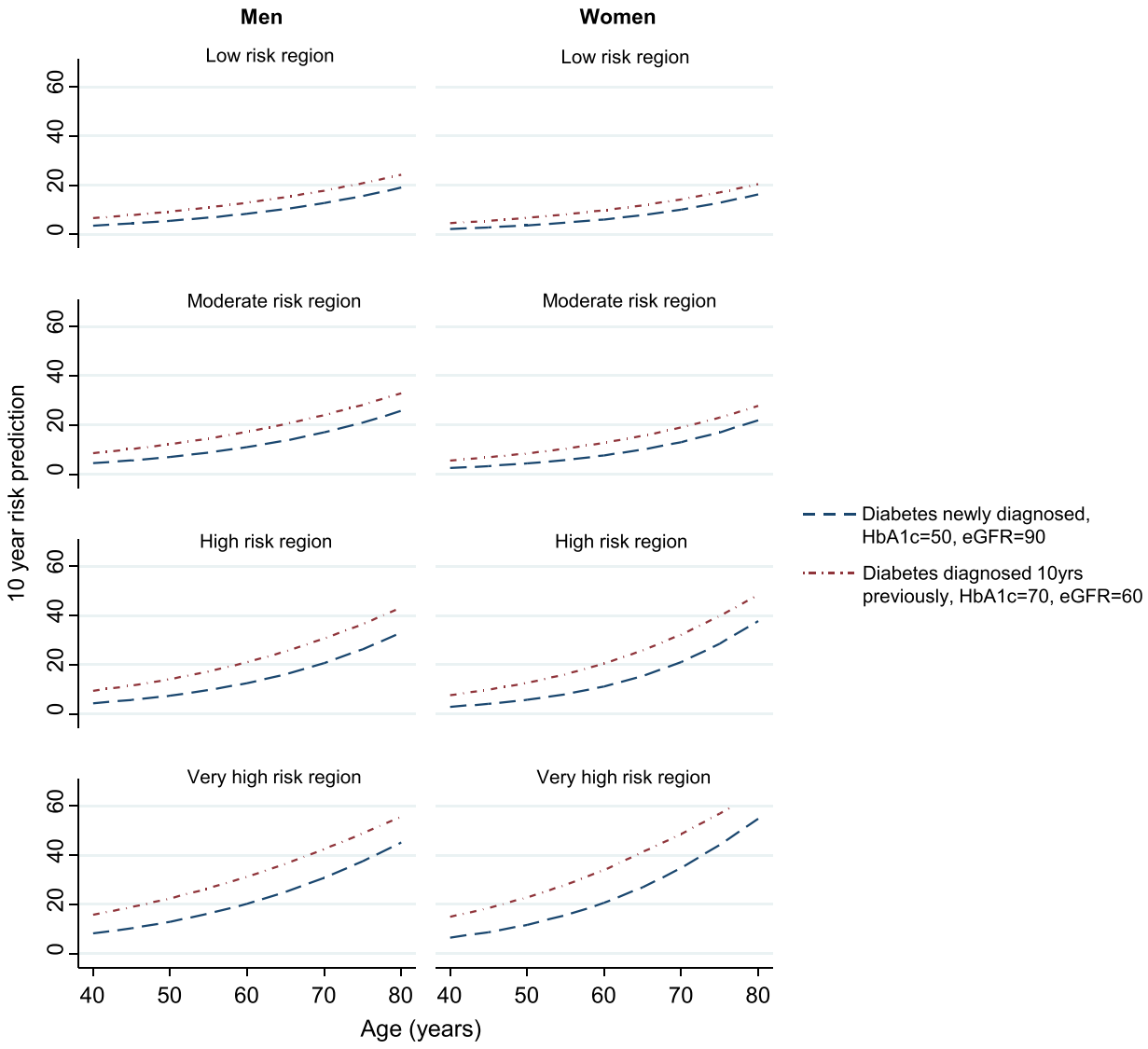


Figure 3 Estimates are for non-smokers with systolic blood pressure of 140 mm Hg, total cholesterol of 5.5 mmol/L and HDL-cholesterol of 1.3 mmol/L. eGFR: estimated Glomerular Filtration Rate (mL/min/1.73 m^2) calculated using the CKD-EPI 2009 equations; HbA1c (mmol/mol): glycated haemoglobin, in International Federation of Clinical Chemistry and Laboratory Medicine (IFCC) units.

Third, while the recalibration applied accounts for substantial variation in whole population levels of risk across Europe, SCORE2-Diabetes also shows good ability to discriminate and provide individual risk estimates for individuals with type 2 diabetes, taking into account their specific risk factors such as age of diabetes diagnosis, HbA1c and kidney function (*Structured Graphical Abstract*). SCORE2-Diabetes can therefore be used to help guide clinicians and individual patients for considering the intensity of existing treatment (such as lipid lowering therapies) as well as additional interventions to prevent CVD (such as sodium-glucose co-transporter 2 inhibitors or glucagon like peptide-1 receptor agonists).

Fourth, development, calibration, validation, and illustration of the SCORE2-Diabetes models have been underpinned by powerful, extensive and complementary datasets of contemporary relevance to individuals with type 2 diabetes across European populations. In particular, SCORE2-Diabetes was developed using data from over 220 000 individuals with type 2 diabetes from 10 different data sources, enhancing the reliable and accurate estimation of risk ratios, and validated using additional data on over 210 000 individuals, ensuring the generalisability and validity of the approach.

Fifth, the approach used in SCORE2-Diabetes accounts for the impact of the competing risk of non-CVD death. This statistical adjustment should prevent any overestimation of CVD risk, thereby reducing the chances of over-estimating the potential benefits of CVD risk-modifying treatments. This adjustment particularly benefits treatment decisions in older individuals, and those from high or very-high-risk regions, where the risk of competing non-CVD deaths is high.

Finally, our analysis has illustrated the performance of SCORE2-Diabetes with simulated data on individuals with type 2 diabetes from different European risk regions, showing that the proportions of individuals across different risk categories are strikingly different across regions. This finding suggests that our risk estimates should assist

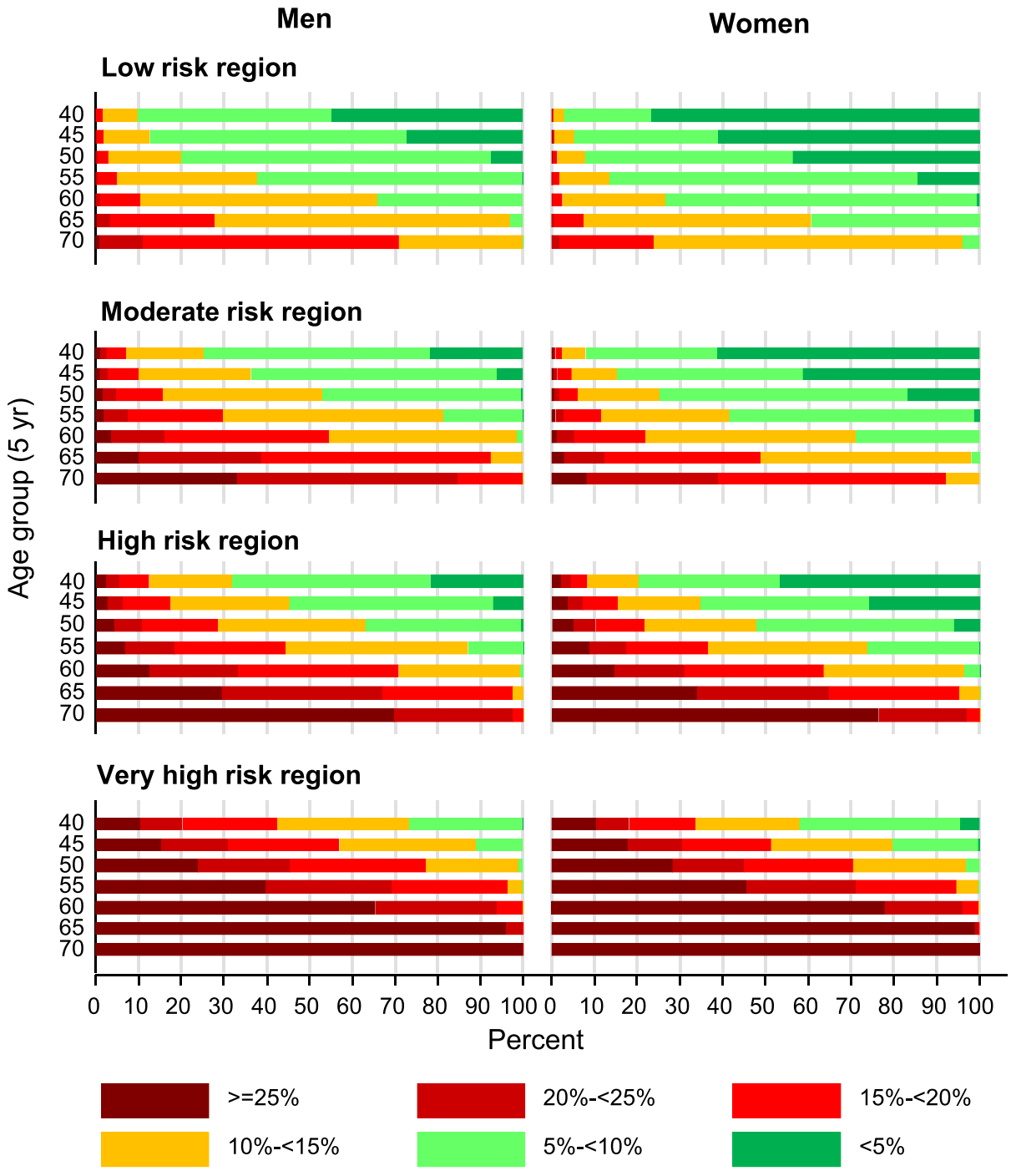


Figure 4 The proportion of individuals expected in each risk category was estimated to reflect the age-group and sex-specific risk factor values in each risk region (see Supplementary data online, *Supplementary methods*).

policy makers to make more appropriate and locally informed decisions about the allocation of resources.

The potential limitations of this study merit consideration. We extended the SCORE2 risk prediction models by estimating additional relative risks for the diabetes-related variables using data sources from European regions and populations at low or moderate CVD risk. Ideally, relative risk estimation for use in high and very high-risk countries would have involved large nationally representative, prospective cohorts in these countries, coupled with prolonged follow-up and validation of fatal and non-fatal CVD endpoints. Unfortunately, such data do not yet exist. Indeed, even in low- and moderate-risk regions, the data sources involved may not be nationally or regionally representative, reflecting past periods of time, 'healthy' volunteers contributing to cohort studies, or, in the case of registry data, individuals

with increased tendency to seek medical attention. However, while such biases can lead to misleading levels of absolute risk, relative risks are generally unaffected.[2,33,34] Furthermore, our analyses identified little heterogeneity in model coefficients across studies used in model derivation, suggesting transferability of model coefficients across different populations, as evidenced by good discrimination in all populations tested. Crucially, SCORE2-Diabetes models were recalibrated using nationally representative incidence rates, an important step not commonly considered in development of other CVD risk scores for individuals with diabetes,[10,11] avoiding the limitations of mis-calibration provided by potentially non-representative incidence rates in derivation datasets.

The rescaling factors used in recalibration of SCORE2-Diabetes were identical to those used in recalibration of the SCORE2 risk models. This approach assumes that the additional measurement of diabetes age at diagnosis, HbA1c, and eGFR among individuals with diabetes does not importantly change the average sex- and age-specific risk predictions for the regional target population (including those with and without diabetes). We have tested this assumption using several datasets mostly from the low and moderate-risk regions, but further testing should be completed if the relevant data become available in the future. Likewise, more accurate representation of the potential predicted risk distributions in each European risk region could be achieved by applying the recalibrated SCORE2-Diabetes models to risk factor levels from the diabetes-specific populations from additional representative datasets in each risk region. In parallel to the analysis presented in this study, we have developed methods and statistical codes that will allow future validation and illustration of SCORE2-Diabetes in diabetes-specific registries as data becomes available.[21,22]

Information on incident non-fatal HF and PAD were not uniformly recorded in available data sources, and therefore it has not been possible to include them in the SCORE2-Diabetes endpoint. However, sensitivity analyses suggested that while discrimination of SCORE2-Diabetes for these outcomes is still likely to be good, estimates of CVD risk could be conservative and may underestimate the potential benefits of CVD risk-modifying treatments that also reduce HF risk. It is assumed that many individuals using SCORE2-Diabetes will already be taking medication that affects CVD risk, and this assumption is respected by inclusion of such individuals in datasets used to derive and recalibrate the models. In addition, some individuals in our model derivation cohorts may have initiated preventative treatment (e.g. statin) during follow-up and accounting for this could improve model calibration and discrimination. However previous analyses have suggested that inclusion of information on statin initiation during follow-up provides only limited improvement in risk prediction.[35] Furthermore, comprehensive individual-participant data on medication use were unavailable in all data sources used for model development and recalibration. This was also the case with family history of CVD, socio-economic status, ethnicity, and albuminuria meaning interpretation of SCORE2-Diabetes estimates may require clinical judgement, especially for individuals for whom these factors may be relevant (e.g. those with a family history of premature CVD, or in higher-risk socio-economic and non-white ethnic groups) as well as in older age groups (those aged over 70 years) where additional consideration of multimorbidities and life expectancy may be needed.[9,36] While the SCORE2-Diabetes models are broadly applicable across all European countries, there remains a place for country-specific risk calculators that consider the specific characteristics relevant to that population (ideally incorporating information on socio-economic status and ethnicity). More generally, better quality data collection, both in terms of risk factors and outcomes, will serve to improve the quality of risk prediction, and should be integral to the evolution of electronic health records and their linkage.

We compared the performance of SCORE2-Diabetes with the ADVANCE model in the SNDR dataset since this dataset is considered nationally representative of the diabetes population in Sweden. However, due to lack of data availability albuminuria was used as a binary rather than continuous variable and atrial fibrillation was not included, meaning that the full predictive ability of ADVANCE may not have been observed in the current analysis. Comparison with other risk models already developed for use in individuals with type 2 diabetes was generally not possible because these models contain variables often not available in datasets. Similarly, data availability for recalibration is very limited, making such models less appropriate for use across different risk regions. The discrimination ability reported for SCORE2-Diabetes, was somewhat lower than that previously reported for CVD risk scores in the general population.[4,5,37,38] However, higher C-indices are expected to be seen in general population cohorts given the wider age range and heterogeneity in risk factor values. Furthermore, general population CVD risk scores are commonly designed for use in individuals both with and without diabetes, with reported C-indices including the considerable discriminative ability of diabetes status itself. The C-indices from the current analysis are similar to those previously reported specifically among individuals with type 2 diabetes.[7] By contrast, the clinical performance of risk prediction models depends importantly on differing ability to predict the correct level of risk in the target population (i.e. extent of 'calibration').[37] We, therefore, ensured SCORE2-Diabetes was well-calibrated to current absolute risk levels for each European region.

In summary, we have derived, recalibrated, validated, and illustrated SCORE2-Diabetes, a 10-year risk model tailored to individuals with diabetes in European populations to predict 10-year risk of first-onset CVD (*Structured Graphical Abstract*). This will assist future guidelines on CVD prevention in individuals with type 2 diabetes, by providing an appropriate risk estimation system to enhance the accuracy, practicability, and sustainability of CVD prevention strategies and help guide preventative treatment.

Acknowledgements

We acknowledge with gratitude the people with diabetes, NHS staff and organisations (the Scottish Care Information-Diabetes Steering Group, the Scottish Diabetes Group, the Scottish Diabetes Survey Group, the diabetes managed clinical networks) involved in providing data, setting up, maintaining and overseeing collation of data for people with diabetes in Scotland. Data linkage was performed by Public Health Scotland. The Scottish Diabetes Research Network is supported by National Health Service (NHS) Research Scotland, a partnership involving Scottish NHS Boards and the Chief Scientist Office of the Scottish Government

We thank investigators and participants of the several studies that contributed data to the Emerging Risk Factors Collaboration (ERFC).

This research has been conducted using the UK Biobank Resource under Application Number 13784.

Data from the Clinical Practice Research Datalink (CPRD) were obtained under licence from the UK Medicines and Healthcare products Regulatory Agency (protocol 162RMn2). CPRD uses data provided by patients and collected by the NHS as part of their care and support.

Supplementary data

Supplementary data is available at *European Heart Journal* online.

Ethical approval

Relevant ethical approval and participant consent were already obtained in all studies that contributed data to this work (Supplementary data online).

Data availability

Data used for the current study are available upon reasonable request and approval of the individual cohorts or collaborative groups. The Scottish data used in this paper can be requested from the Public Benefit and Privacy Panel for Health and Social Care—more details are available from: https://www.informationgovernance.scot.nhs.uk/pbpphsc/.

Conflict of interest

A.W. reports institutional grants from an Alan Turing Institute, British Heart Foundation, NIHR, and Innovative Medicines Initiative. A.T. reports receiving payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events made to him from European Heart Journal: Quality of Care and Clinical Outcomes for being an Editor, reports leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid, for being a member of Trial Steering Committees for SCOT HEART 2 and DUAL ACS 2, and reports owning two shares in two start-up companies called HD Clinical and RE-COGNITION Health Ltd. B.A.F. reports grants or contracts from Novartis, Amgen, Pfizer and Esperion, consulting fees from Novartis, Amgen, Regeneron, Eli Lilly, Sanofi, Merck, AstraZeneca, Pfizer, Viatris, Novo Nordisk, New Amsterdam, Mylan, Daiichi Sankyo, Ionis Pharmaceuticals, dalCOR, CiVi Pharmaceuticals, KrKa Pharmaceuticals, American College of Cardiology, European Atherosclerosis Society and European Society of Cardiology, payments or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Novartis, Amgen, Regeneron, Eli Lilly, Sanofi, Merck, AstraZeneca, Pfizer, Viatris, Novo, Nordisk, New Amsterdam, Mylan, Daiichi Sankyo, Ionis Pharmaceuticals and dalCOR, patents planned, issued or pending from Novartis, Amgen, Regeneron, Sanofi, Merck, AstraZeneca, Novo Nordisk and New Amsterdam, and participation on a Data Safety Monitoring Board or Advisory Board for Novartis, Amgen, Regeneron, Sanofi, Merck, AstraZeneca, Novo Nordisk and New Amsterdam. B.E. reports institutional unrestricted research support to institution from Konung Gustaf V:s och Drottning Victorias Frimurarestiftelse, consulting fees from Novo Nordisk, Sanofi, Eli Lilly, Boehringer Ingelheim, Mundipharma, AMGEN, AstraZeneca and MSD, and payments or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Novo Nordisk, Sanofi, Eli Lilly, Boehringer Ingelheim, Mundipharma, AstraZeneca and MSD.

C.P. reports grants or contracts from MRC DTP PhD Studentship. C.SN. reports support for present manuscript from German Federal Ministry of Health through providing funding to conduct national health surveys BGS98 and DEGS1, and grants or contracts from Federal Ministry of Health (Germany). D.M. reports consulting fees from Ferrer, Merch, Sharp & Dohme, Novo Nordisk and Sanofi, and receiving payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Almirall, Amgen, Esteve, Ferrer, Jansen, Lilly, Merch, Sharp & Dohme, Novo Nordisk and Sanofi. D.V. reports receiving grants or contracts from Bayer, Sanofi, Aventis, Novo Nordisk and Boehringer Ingelheim, participation on a Data Safety Monitoring Board or Advisory Board for FENO, leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid, for European Diabetes Epidemiology Group as a Chair of the Steering Committee, and reports owning stock or stock options at Novo Nordisk. E.D.A. reports support from British Heart Foundation and NIHR Senior Investigator Award, and leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid from Our Future Health, EURAC. F.B. received payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing, or educational events from Bayer, and support for attending meetings and/or travel from Bayer and Novo Nordisk. J.D. reports support from British Heart Foundation Professorship and NIHR Senior Investigator Award, grants, personal fees and non-financial support from Merck Sharp & Dohme, Novartis, Pfizer, and AstraZeneca outside the submitted work, royalties or licences from Wellcome Trust Sanger Institute, leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid from Cambridge University Hospital NHS Foundation Trust, UK Biobank, MRC International Advisory Group, MRC High Throughput Science 'Omics Panel, Scientific Advisory Committee for Sanofi, Wellcome Sanger Institute, Novartis, and Nightingale Health. J.FN. reports consulting fees from Novo Nordisk and Boehringuer Ingelheim and unpaid leadership or fiduciary role in other board, society, committee or advocacy group from Fundación redGDPS. L.Po. reports receiving payments or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Sanofi, Lilly, Boehringer Ingelheim, AstraZeneca and Novo Nordisk, and participation on a Data Safety Monitoring Board or Advisory Board for Novo Nordisk. M.H. reports consulting fees from Medical Park Cardiac Rehab Group and Abbott, payment or honoraria for lectures, presentations, speakers bureaus, manuscript from BMS, Amgen, Daichii-Sankyo, and Medi, leadership or fiduciary role in other board, society, committee or advocacy group, paid or unpaid from European Association of Preventive Cardiology. N.H. reports support for present manuscript from NHS England. N.S. reports grants paid to institution from AstraZeneca, Boehringer Ingelheim, Novartis and Roche Diagnostics, consulting fees from Abbott Laboratories, Afimmune, Amgen, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Hanmi Pharmaceuticals, Merck, Shatp & Dohme, Novartis, Novo Nordisk, Pfizer, Roche Diagnostics and Sanofi, and payment or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Abbott Laboratories, AstraZeneca, Boehringer Ingelheim, Eli Lilly, Janssen and Novo Nordisk. O.T. reports receiving grants or contracts paid to institution and to himself from British Heart Foundation Centre of Research Excellence PhD Stipend RHAG/372. P.R. reports grants or contracts from AstraZeneca, Novo Nordisk and Bayer AG, consulting fees from Astellas, AstraZeneca, Bayer, Boehringer Ingelheim, Eli Lilly, Merck, Novo Nordisk, Sanofi, Aventis and Gilead. P.V. reports consulting fees from Hygeia Hospitals Group and European Society of Cardiology. T.P. reports grants or contracts from Sanofi, and payments or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events from Sanofi and Salveo. R.G. reports payments or honoraria for lectures, presentations, speakers bureaus, manuscript writing or educational events made to him from Annals of Epidemiology due to being an Associate Editor. S.Ka. reports support for present manuscript from British Heart Foundation, UK Medical Research Council and National Institute for Health and Care Research. S.Ki. reports receiving grants or contract from VASCage (Research Centre on Vascular Ageing and Stroke, project number 868624) of the Austrian Research Promotion Agency FFG (COMET program–Competence Centers for Excellent Technologies) funded by the Federal Ministry for Climate Protection, Environment, Energy, Transport, Innovation and Technology; the Federal Ministry for Labour and Economy; and the federal states Tyrol (via Standortagentur), Salzburg, and Vienna (via Vienna Business Agency). W.H. reports support for present manuscript from Medical Research Council UK and Kidney Research UK, grants or contracts from Boehringer Ingelheim and Eli Lilly, and participation on a Data Safety Monitoring Board or Advisory Board for Bayer. Z.X. reports grants or contracts from China Scholarship Council.

Funding

The ERFC co-ordinating centre was underpinned by programme grants from the British Heart Foundation (SP/09/002; RG/13/13/30194; RG/18/13/33946), BHF Centre of Research Excellence (RE/18/1/34212), the UK

Medical Research Council (MR/L003120/1), and the National Institute for Health and Care Research (NIHR) Cambridge Biomedical Research Centre (BRC-1215–20014; NIHR203312), with project-specific support received from the UK NIHR [*], British United Provident Association (BUPA) Foundation and an unrestricted educational grant from GlaxoSmithKline. A variety of funding sources have supported recruitment, follow-up, and laboratory measurements in the studies contributing data to the ERFC, which are listed on the ERFC website (www.phpc.cam.ac.uk/ceu/erfc/list-of-studies).

*The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

This work was supported by Health Data Research UK, which is funded by the UK Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, Department of Health and Social Care (England), Chief Scientist Office of the Scottish Government Health and Social Care Directorates, Health and Social Care Research and Development Division (Welsh Government), Public Health Agency (Northern Ireland), British Heart Foundation, and Wellcome.

Appendix

SCORE2-Diabetes Working Group and the ESC Cardiovascular Risk Collaboration

Lisa Pennells* (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Stephen Kaptoge* (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Helena Bleken Østergaard* (Department of Vascular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands); Stephanie H Read* (Usher Institute, University of Edinburgh, Edinburgh, UK); Fabrizio Carinci* (Department of Statistical Sciences, University of Bologna, Italy); Josep Franch-Nadal (Center for Biomedical Research on Diabetes and Associated Metabolic Diseases (CIBERDEM), Instituto de Salud Carlos III (ISCIII), Barcelona, Spain. Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gorina (IDIAPJGol), Barcelona, Spain. Primary Health Care Center Raval Sud, Institut Català de la Salut, Barcelona, Spain); Carmen Petitjean (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Owen Taylor (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Steven H.J. Hageman (Department of Vascular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands); Zhe Xu (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Fanchao Shi (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Sarah Spackman (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Stefano Gualdi (Internet Express, Pescara, Italy); Naomi Holman (Department of Epidemiology and Biostatistics, Imperial College London, UK); Rui Bebiano Da Providencia E Costa (UCL Institute for Health Informatics Research, University College London, London, UK); Fabrice Bonnet (University of Rennes 1, Rennes CHU, Rennes, France); Hermann Brenner (Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany. Network Aging Research, University of Heidelberg, Heidelberg, Germany); Richard F Gillum (Howard University Hospital, Washington DC, DC, USA); Stefan Kiechl (Department of Neurology, Medical University Innsbruck, Innsbruck, Austria); Deborah A Lawlor (MRC Integrative Epidemiology Unit at the University of Bristol, UK; Population Health Science, Bristol Medical School, Bristol University, UK); Louis Potier (Université Paris City, Bichat Hospital, AP-HP, INSERM U1151 INEM, Paris, France); Ben Schöttker (Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany. Network Aging Research, University of Heidelberg, Heidelberg, Germany); Reecha Sofat (UCL Institute for Health Informatics Research, University College London, London, UK); Henry Völzke (Institute for Community Medicine, University Medicine Greifswald, Germany); Johann Willeit (Department of Neurology, Medical University Innsbruck, Innsbruck, Austria); Zane Baltane (The Centre for Disease Prevention and Control of Latvia); Stephen Fava (University of Malta); Sandor Janos (University of Debrecen); Astrid Lavens (Sciensano, Belgium); Santa Pildava (The Centre for Disease Prevention and Control of Latvia); Tamara Poljicanin MD, PhD (Croatian Institute of Public Health, Rockefellerova 7, Zagreb, Croatia); Ivan Pristas MD, PhD (Croatian Institute of Public Health, Rockefellerova 7, Zagreb, Croatia); Peter Rossing (Steno Diabetes Center, Gentofte, Denmark); Reiff Sascha (Ministry of Health, Malta); Christa Scheidt-Nave (Robert Koch Institute); Iztok Stotl (Department of Endocrinology, Diabetes and Metabolic Diseases,University Medical Centre Ljubljana, Ljubljana, Slovenia); Gail Tibor (University of Debrecen); Vilma Urbančič-Rovan (Department of Endocrinology, Diabetes and Metabolic Diseases,University Medical Centre Ljubljana, Ljubljana, Slovenia Faculty of Medicine, University of Ljubljana, Ljubljana, Slovenia); An-Sofie Vanherwegen (Sciensano, Belgium); Dorte Vistisen (Steno Diabetes Center, Gentofte, Denmark); Yong Du (Robert Koch Institute); Matthew R Walker (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Peter Willeit (Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Innsbruck, Austria, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Brian Ference (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Dirk De Bacquer (Ghent University, Ghent, Belgium); Martin Halle (University Hospital ´Klinikum rechts der Isar´, Technical University of Munich, Munich, Germany); Radu Huculeci (European Society of Cardiology, Brussels, Belgium); John William McEvoy (National University of Ireland Galway, Galway, Ireland); Adam Timmis (William Harvey Research Institute, Barts & The London School of Medicine & Dentistry, Queen Mary University of London, London, UK); Panagiotis Vardas (Heraklion University Hospital, Crete, Greece); Jannick A. N. Dorresteijn (Department of Vascular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands); Ian Graham (School of Medicine, Trinity College Dublin, The University of Dublin, College Green, Dublin, Ireland); Angela Wood (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Björn Eliasson (Department of Medicine, Sahlgrenska University Hospital, Gothenburg, Sweden; The Swedish National Diabetes Register, Gothenburg, Sweden); William Herrington (Medical Research Council Population Health Research Unit at the University of Oxford, Nuffield Department of Population Health, University of Oxford, Oxford, UK); John Danesh (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK); Dídac Mauricio (Department of Endocrinology & Nutrition, Hospital de la Santa Creu i Sant Pau, and Sant Pau Biomedical Research Institute, Barcelona, Spain. Center for Biomedical Research on Diabetes and Associated Metabolic Diseases (CIBERDEM), Instituto de Salud Carlos III (ISCIII), Barcelona, Spain. Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gorina (IDIAPJGol), Barcelona, Spain. Departament de Medicine, Central University of Catalonia, Vic, Spain); Massimo Massi Benedetti†

(Hub for International Health Research, Perugia,Italy); Naveed Sattar† (Institute of Cardiovascular & Medical Sciences, University of Glasgow, Glasgow, UK); Frank L.J. Visseren† (Department of Vascular Medicine, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands); Sarah Wild† (Usher Institute, University of Edinburgh, Edinburgh, UK); Emanuele Di Angelantonio† ‡ (Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK, Health Data Science Centre, Human Technopole, Milan, Italy).

*Contributed equally.

†Contributed equally.

‡Corresponding author. email: ed303@medschl.cam.ac.uk

Investigators from contributing data sources

ERFC Investigators:

DESIR: Beverley Balkau, Fabrice Bonnet, Frederic Fumeron; ESTHER: Hannah Stocker, Bernd Holleczek; SHIP: Sabine Schipf, Carsten Oliver Schmidt, Marcus Dörr; BRUN: Herbert Tilg, Christoph Leitner, Marlene Notdurfter; BWHHS: Julie Taylor, Caroline Dale, David Prieto-Merino; NHANES3: Richard F Gillum

EUBIROD Investigators:

BELGIUM: Astrid Lavens, An-Sofie Vanherwegen; CROATIA: Tamara Poljicanin, Ivan Pristas, Tamara Buble, Pero Ivanko; DENMARK: Peter Rossing, Bendix Carstensen; GERMANY: Christin Heidemann, Yong Du, Christa Scheidt-Nave; HUNGARY: Tibor Gall, Janos Sandor; LATVIA: Zane Baltane, Santa Pildava, Jana Lepiksone; MALTA: Caroline J. Magri, Joseph Azzopardi; SLOVENIA: Iztok Stotl.

SIDIAP Investigators: Jordi Real, Bogdan Vlacho, Manel Mata-Cases.

References

1. Timmis A, Vardas P, Townsend N, Torbica A, Katus H, De Smedt D, *et al.* European Society of cardiology: cardiovascular disease statistics 2021. *Eur Heart J* 2022;**43**:716–799. https://doi.org/10.1093/eurheartj/ehab892
2. Emerging Risk Factors Collaboration. Diabetes mellitus, fasting blood glucose concentration, and risk of vascular disease: a collaborative meta-analysis of 102 prospective studies. *Lancet* 2010;**375**:2215–2222. https://doi.org/10.1016/S0140-6736(10)60484-9
3. Visseren FLJ, Mach F, Smulders YM, Carballo D, Koskinas KC, Bäck M, *et al.* 2021 ESC guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J*. 2021;**42**:3227–3337. https://doi.org/10.1093/eurheartj/ehab484
4. Goff DC, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, *et al.* 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Circulation* 2014;**129**:S49–S73. https://doi.org/10.1161/01.cir.0000437741.48606.98
5. WHO CVD Risk Chart Working Group. World health organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. *Lancet Glob Health* 2019;**7**:e1332–e1345. https://doi.org/10.1016/S2214-109X(19)30318-3
6. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* 2017;**357**; j2099. https://doi.org/10.1136/bmj.j2099
7. Dziopa K, Asselbergs FW, Gratton J, Chaturvedi N, Schmidt AF. Cardiovascular risk prediction in type 2 diabetes: a comparison of 22 risk scores in primary care settings. *Diabetologia* 2022;**65**:644–656. https://doi.org/10.1007/s00125-021-05640-y
8. Read SH, van Diepen M, Colhoun HM, Halbesma N, Lindsay RS, McKnight JA, *et al.* 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
9. Berkelmans GFN, Gudbjornsdottir S, Visseren FLJ, Wild SH, Franzen S, Chalmers J, *et al.* Prediction of individual life-years gained without cardiovascular events from lipid, blood pressure, glucose, and aspirin treatment based on data of more than 500 000 patients with type 2 diabetes mellitus. *Eur Heart J*. 2019;**40**:2899–2906. https://doi.org/10.1093/eurheartj/ehy839
10. Kengne AP, Patel A, Marre M, Travert F, Lievre M, Zoungas S, *et al.* Contemporary model for cardiovascular risk prediction in people with type 2 diabetes. *Eur J Cardiovasc Prev Rehabil*. 2011;**18**:393–398. https://doi.org/10.1177/1741826710394270
11. Stevens RJ, Kothari V, Adler AI, Stratton IM; United Kingdom Prospective Diabetes Study Group. The UKPDS risk engine: a model for the risk of coronary heart disease in type II diabetes (UKPDS 56). *Clin Sci (Lond)* 2001;**101**:671–679. https://doi.org/10.1042/CS20000335
12. Hageman S, Pennells L, Ojeda F, Kaptoge S, Kuulasmaa K, de Vries T, *et al.* SCORE2 Risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. *Eur Heart J*. 2021;**42**:2439–2454. https://doi.org/10.1093/eurheartj/ehab309
13. McKnight JA, Morris AD, Cline D, Peden N, Fischbacher C, Wild S. Implementing a national quality assurance system for diabetes care: the Scottish diabetes survey 2001–2006. *Diabet Med* 2008;**25**:743–746. https://doi.org/10.1111/j.1464-5491.2008.02453.x
14. Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, *et al.* Data resource profile: clinical practice research datalink (CPRD). *Int J Epidemiol*. 2015;**44**:827–836. https://doi.org/10.1093/ije/dyv098
15. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, *et al.* UK Biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS Med*. 2015;**12**:e1001779. https://doi.org/10.1371/journal.pmed.1001779
16. Emerging Risk Factors Collaboration, Danesh J, Erqou S, Walker M, Thompson SG, Tipping R, Ford C, *et al.* The emerging risk factors collaboration: analysis of individual data on lipid, inflammatory and other markers in over 1.1 million participants in 104 prospective studies of cardiovascular diseases. *Eur J Epidemiol*. 2007;**22**:839–869. https://doi.org/10.1007/s10654-007-9165-7
17. Emerging Risk Factors Collaboration, Angelantonio E, Kaptoge S, Wormser D, Willeit P, Butterworth AS, Bansal N, *et al.* Association of cardiometabolic multimorbidity with mortality. *JAMA* 2015;**314**:52–60. https://doi.org/10.1001/jama.2015.7008
18. Gudbjornsdottir S, Cederholm J, Nilsson PM, Eliasson B; Steering Committee of the Swedish National Diabetes Register. The national diabetes register in Sweden: an implementation of the st. Vincent declaration for quality improvement in diabetes care. *Diabetes Care*. 2003;**26**:1270–1276. https://doi.org/10.2337/diacare.26.4.1270
19. Mata-Cases M, Mauricio D, Real J, Bolibar B, Franch-Nadal J. Is diabetes mellitus correctly registered and classified in primary care? A population-based study in Catalonia, Spain. *Endocrinol Nutr*. 2016;**63**:440–448. https://doi.org/10.1016/j.endonu.2016.07.004
20. Bolibar B, Fina Aviles F, Morros R, Garcia-Gil Mdel M, Hermosilla E, Ramos R, *et al.* SIDIAP database: electronic clinical records in primary care as a source of information for epidemiologic research. *Med Clin (Barc)* 2012;**138**:617–621. https://doi.org/10.1016/j.medcli.2012.01.020
21. Carinci F, Štotl I, Cunningham SG, Poljicanin T, Pristas I, Traynor V, *et al.* Making use of comparable health data to improve quality of care and outcomes in diabetes: the EUBIROD review of diabetes registries and data sources in Europe. *Front Clin Diabetes Healthc*. 2021;**2**:744516. https://doi.org/10.3389/fcdhc.2021.744516
22. Cunningham SG, Carinci F, Brillante M, Leese GP, McAlpine RR, Azzopardi J, *et al.* Core standards of the EUBIROD project. Defining a European diabetes data dictionary for clinical audit and healthcare delivery. *Methods Inf Med* 2016;**55**:166–176. https://doi.org/10.3414/ME15-01-0016
23. EUBIROD. NeuBIRO Software http://www.eubirod.eu/academy/software.html. Accessed December 2022.
24. Di Iorio CT, Carinci F, Oderkirk J, Smith D, Siano M, de Marco DA, *et al.* Assessing data protection and governance in health information systems: a novel methodology of privacy and ethics impact and performance assessment (PEIPA). *J Med Ethics* 2021;**47**:e23. https://doi.org/10.1136/medethics-2019-105948
25. Di Iorio CT, Carinci F, Brillante M, Azzopardi J, Beck P, Bratina N, *et al.* Cross-border flow of health information: is 'privacy by design' enough? Privacy performance assessment in EUBIROD. *Eur J Public Health* 2013;**23**:247–253. https://doi.org/10.1093/eurpub/cks043
26. Di Iorio CT, Carinci F, Azzopardi J, Baglioni V, Beck P, Cunningham S, *et al.* Privacy impact assessment in the design of transnational public health information systems: the BIRO project. *J Med Ethics*. 2009;**35**:753–761. https://doi.org/10.1136/jme.2009.029918
27. Holman N, Knighton P, Wild SH, Sattar N, Dew C, Gregg EW, *et al.* Cohort profile: national diabetes audit for England and Wales. *Diabet Med* 2021;**38**:e14616. https://doi.org/10.1111/dme.14616
28. Wolbers M, Koller MT, Witteman JC, Steyerberg EW. Prognostic models with competing risks: methods and application to coronary risk prediction. *Epidemiology*. 2009;**20**:555–561. https://doi.org/10.1097/EDE.0b013e3181a39056
29. Inker LA, Eneanya ND, Coresh J, Tighiouart H, Wang D, Sang Y, *et al.* New creatinine- and cystatin C-based equations to estimate GFR without race. *N Engl J Med* 2021;**385**:1737–1749. https://doi.org/10.1056/NEJMoa2102953
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. *Ann Intern Med* 2015;**162**:55–63. https://doi.org/10.7326/M14-0697
31. European Society of Cardiology. https://www.escardio.org/Education/ESC-Prevention-of-CVD-Programme/Risk-assessment/esc-cvd-risk-calculation-app. Accessed December 2022.
32. European Society of Cardiology. https://www.escardio.org/Education/Practice-Tools/CVD-prevention-toolbox/HeartScore. Accessed December 2022.
33. Batty GD, Gale CR, Kivimaki M, Deary IJ, Bell S. Comparison of risk factor associations in UK biobank against representative, general population based studies with conventional response rates: prospective cohort study and individual participant meta-analysis. *BMJ*. 2020;**368**:m131. https://doi.org/10.1136/bmj.m131
34. Matsushita K, Coresh J, Sang Y, Chalmers J, Fox C, Guallar E, *et al.* Estimated glomerular filtration rate and albuminuria for prediction of cardiovascular outcomes: a collaborative meta-analysis of individual participant data. *Lancet Diabetes Endocrinol* 2015;**3**:514–525. https://doi.org/10.1016/S2213-8587(15)00040-6

35. Xu Z, Arnold M, Stevens D, Kaptoge S, Pennells L, Sweeting MJ, *et al.* Prediction of cardiovascular disease risk accounting for future initiation of statin treatment. *Am J Epidemiol* 2021;**190**:2000–2014. https://doi.org/10.1093/aje/kwab031
36. de Vries T, Cooney MT, Selmer, RM, Hageman SHJ, Pennells LA, Wood A, *et al.* SCORE2-OP Risk prediction algorithms: estimating incident cardiovascular event risk in older persons in four geographical risk regions. *Eur Heart J* 2021;**42**:2455–2467. https://doi.org/10.1093/eurheartj/ehab312
37. Pennells L, Kaptoge S, Wood A, Sweeting M, Zhao X, White I, *et al.* Equalization of four cardiovascular risk algorithms after systematic recalibration: individual-participant meta-analysis of 86 prospective studies. *Eur Heart J* 2019;**40**:621–631. https://doi.org/10.1093/eurheartj/ehy653
38. Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, *et al.* Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;**336**:1475–1482. https://doi.org/10.1136/bmj.39609.449676.25