Prediction of individual lifetime cardiovascular risk and potential treatment benefit: development and recalibration of the LIFE-CVD2 model to four European risk regions

LIFE-CVD2 Working Group*

*authors and investigators are listed at the end of the manuscript

Abstract

Aim: The 2021 European Society of Cardiology prevention guidelines recommend the use of (lifetime) risk prediction models to aid decisions regarding initiation of prevention. We aimed to update and systematically recalibrate the LIFEtime-perspective CardioVascular Disease (LIFE-CVD) model to four European risk regions for the estimation of lifetime CVD risk for apparently healthy individuals.

Methods and Results: The updated LIFE-CVD (i.e., LIFE-CVD2) models were derived using individual-participant data from 44 cohorts in 13 countries (687,135 individuals without established CVD, 30,939 CVD events in median 10.7 years of follow-up). LIFE-CVD2 uses sex-specific functions to estimate the lifetime risk of fatal and non-fatal CVD events with adjustment for the competing risk of non-CVD death, and is systematically recalibrated to four distinct European risk regions. The updated models showed good discrimination in external validation among 1,657,707 individuals (61,311 CVD events) from eight additional European cohorts in seven countries, with a pooled C-index of 0.795 (95%CI 0.767-0.822). Predicted and observed CVD event risks were well calibrated in population-wide electronic health records data in the UK (CPRD) and Netherlands (ELAN). When using LIFE-CVD2 to estimate potential gain in CVD-free life expectancy from preventive therapy, projections varied by risk region reflecting important regional differences in absolute lifetime risk. For example a 50-year-old smoking woman with a SBP of 140 mm Hg was estimated to gain 0.9 years in the low risk region versus 1.6 years in the very high risk region from lifelong 10 mm Hg SBP reduction. The benefit of smoking cessation for this individual ranged from 3.6 years in the low risk region to 4.8 years in the very high risk region.

Interpretation: By taking into account geographical differences in CVD incidence using contemporary representative data sources, the recalibrated LIFE-CVD2 model provides a more accurate tool for the prediction of lifetime risk and CVD-free life expectancy for individuals without previous CVD, facilitating shared decision-making for cardiovascular prevention as recommended by 2021 European guidelines.

2

Lay summary

The study introduces LIFE-CVD2, a new tool that helps predict the risk of heart disease over a person's lifetime, and highlights how where you live in Europe can affect this risk.

- Using health information from over 687,000 people, LIFE-CVD2 looks at things like blood pressure and whether someone smokes to figure out their chance of having heart problems later in life. Health information from another 1.6 million people in seven different European countries was used to show that it did a good job of predicting who might develop heart disease.

- Knowing your heart disease risk over your whole life helps doctors give you the best advice to keep your heart healthy. Let's say there's a 50-year-old woman who smokes and has a bit high blood pressure. Right now, she might not look like she's in danger. But with the LIFE-CVD2 tool, doctors can show her how making changes today, like lowering her blood pressure or stopping smoking, could mean many more years without heart problems. These healthy changes can make a big difference over many years.

Introduction

Cardiovascular diseases (CVD), which include coronary heart disease and stroke, are the most common fatal non-communicable diseases globally, responsible for an estimated 18.6 million deaths in 2019.[1] CVD remains a major cause of morbidity and mortality in Europe. A key strategy in the prevention of CVD is the use of risk prediction algorithms to target preventive interventions on people who benefit from them most.[2,3] Recent European Society of Cardiology (ESC) CVD prevention guidelines recommend an individualized strategy regarding initiation of preventive treatment, in which personal preferences, expected treatment side effects, predicted 10-year CVD risk, and/or lifetime risk estimates are all taken into account in a shared decision making process.[4] Since age is the primary driver of 10-year CVD risk, younger individuals with a high risk factor burden may remain below treatment thresholds based on 10-year risk estimates. However, the potential long-term gain in CVD-free life expectancy from preventive treatment in such individuals may be substantial.[5] Conversely, older individuals often have high 10-year CVD risks, but can have limited treatment benefit due to shorter remaining life expectancy. Use of lifetime risk predictions and associated projected lifetime benefits of preventive therapies can therefore support patient-doctor communication and shared decision making, an approach that is especially recommended for younger people in the 2021 ESC CVD prevention guidelines.[4]

For apparently healthy individuals, the previously published LIFE-CVD model was derived in the MESA study to predict 10-year and lifetime CVD risk.[5] Risk prediction algorithms developed in one population may over- or underestimate risk in another population (i.e. they may not be well 'calibrated'), since CVD event rates and average risk factor levels vary with time and geographic region. Performance can be improved by statistical adjustment (recalibration) to contemporary CVD event rates and risk factor levels.[6] No existing risk prediction model for estimation of lifetime CVD risk has been systematically recalibrated to estimate risk in different regions of Europe.

4

Here we describe the development of the LIFEtime-perspective CardioVascular Disease 2 (LIFE-CVD2) model, which is a major update of the previous LIFE-CVD model in two ways.[5] First, additional datasets from multiple countries are included in model derivation to ensure generalisability; second the models were recalibrated to four European risk regions using contemporary and representative risk factor levels and CVD incidence. The recalibration approach was aligned with similar methods used for 10-year risk CVD risk algorithms currently recommended by international guidelines.[7–9]

Methods

Study design

The updated LIFE-CVD2 model involved multiple data sources for all stages of the project (**Figure 1**). First, to enable reliable estimation of age- and sex-specific relative risks, prediction models for cardiovascular events and non-CVD mortality were derived using individual-participant data from 44 prospective cohorts involving 687,135 participants in 13 countries, including the MESA study in which the original LIFE-CVD model was derived. Second, to adapt risk prediction models to the circumstances of four European risk region (grouped on age- and sex-standardized CVD mortality rates, as with the SCORE2 risk algorithms)[8], the derived risk models were recalibrated using estimated contemporary age- and sex-specific incidences and risk factor distributions. Third, to enhance validity and generalisability, we completed external validation using individual-participant data from a further 8 European prospective cohorts (i.e., studies not used in the model derivation) from 7 countries, involving 1,657,707 participants without prior CVD.

Derivation and internal validation data

The updated LIFE-CVD2 model was derived using individual-participant data from cohorts included in the Emerging Risk Factor Collaboration (ERFC) and the UK Biobank (UKB).[10,11] The ERFC has collated and harmonised individual-participant data from many long-term prospective cohort studies of CVD risk factors and outcomes. Prospective studies in the ERFC

5

were included in this analysis if they met all the following criteria: had recorded baseline information on risk factors necessary to derive risk prediction models (age, sex, current smoking status (versus never of former smoking), history of diabetes mellitus (DM), systolic blood pressure (SBP), and total- and HDL-cholesterol); were approximately population-based (i.e., did not select participants on the basis of 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 1-year of median follow-up. The UKB is a single large prospective cohort study with individual-participant data on approximately 500,000 participants aged >40 years recruited across 23 UK based assessment centres during 2006-2010, and followed-up for cause-specific morbidity and mortality through linkages to routinely available national datasets and disease-specific registers. Individuals with prior CVD at baseline were excluded from model derivation. Details of contributing cohorts are provided in **Supplementary Table 1**.

External validation data

For external validation eight additional data sources were used: the Clinical Practice Research Datalink (CPRD, United Kingdom, low risk region), the Extramural LUMC Academic Network (ELAN, The Netherlands, low risk region)[12], Heinz Nixdorf Recall study (HNR, Germany, moderate risk region)[13], the Estonian Biobank (high risk region)[14] and the Health, Alcohol and Psychosocial factors In Eastern Europe study (HAPIEE, including cohorts from: Poland and Czech Republic [high risk region], Russia and Lithuania [very high risk region]).[15]

The CPRD is a UK based primary care database of anonymized medical records from 674 general practices, with coverage of over 11.3 million patients and is broadly representative of the general population in terms of age, sex and ethnicity. The data used for this study are restricted to the region of England with baseline during the period April 01, 2004 to 2006 and

6

follow-up to November 30, 2017. Incident nonfatal events are obtained from linkage with Hospital Episode Statistics and deaths from the Office for National Statistics. ELAN is a prospective, population-based study using routine health care data from general practitioners with a baseline in 2010, linked to hospital and registry data from Statistics Netherlands, in the region of The Hague and Leiden, the Netherlands. HNR is a population-based study in the large, heavily industrialized Ruhr area, Germany. From December 2000 to August 2003 random samples of men and women aged 45-75 were drawn from mandatory residency lists of three cities in the Ruhr area of North-West Germany. The Estonian Biobank is a population-based biobank of the Estonian Genome Center of the University of Tartu. Follow-up of incident fatal and non-fatal coronary heart disease and stroke events of a subset of the cohort is on-going as our database is being linked with the national healthcare registries and regional and central hospital databases. The HAPIEE study comprises four prospective urban population based cohorts from Eastern Europe, located in Novosibirsk (Russia), Krakow (Poland), Kaunas (Lithuania), and six cities of the Czech Republic. Each cohort recruited a random sample of men and women aged 45-69 years at baseline conducted in 2002–2005 (2005–2008 in Lithuania), stratified by sex and 5-year age groups. From these cohorts, all individuals aged 35 and older without prior CVD or DM were included.

Statistical analysis

The LIFE-CVD2 model comprised complimentary sex-specific Cox proportional hazards models for cardiovascular events and non-CVD mortality respectively. These models used age as the time axis (i.e. left truncation), were stratified by cohort, and included the same predictors as the ESC recommended SCORE2 and SCORE2-OP 10-year risk models,[8,8] namely: age, SBP, total and HDL-cholesterol, current smoking status, and DM and their interactions with age at entry. While the LIFE-CVD2 risk models are not intended for use in individuals with diabetes, participants with a history of diabetes were included at the model derivation stage (with appropriate adjustment for diabetes status), since it wasn't possible to

7

exclude people with diabetes from population-level mortality statistics and risk factor data used in recalibration. Lifetime predictions are generated by calculating individual-participant specific cumulative survival estimates from the CVD and non-CVD models repetitively in one-year windows across all future life years using a life table approach[16] which incorporates adjustment for the competing risk of non-CVD death. This approach has been shown to yield accurate CVD risk predictions beyond the original cohort follow-up for up to 17 year risk predictions.[16] The CVD-free life-expectancy can be read from this life table and is defined as the median survival without a CVD-event or death, the age at which the cumulative survival probability becomes <0.5. The 45-90 year age range of the original LIFE-CVD algorithm has been extended to 35-100 years for LIFE-CVD2 to allow for lifetime predictions for younger individuals and to be able to directly model beyond a survival probability of <0.5 for older individuals. Lifetime risk was defined as the risk of having a CVD event before the age of 80 years life and was calculated by summing all yearly predicted event risks from the current age up until the age of 80 years. Since not all data sources included in the derivation data had follow-up across the complete LIFE-CVD2 age range, the 'un-calibrated' baseline hazard was based upon the low risk region incidence data (WHO CVD mortality rates converted to CVD incidence) to ensure smooth baseline survival curves. The primary outcome predicted by LIFE-CVD2 was defined as a composite of cardiovascular mortality, non-fatal myocardial infarction and non-fatal stroke. The competing non-CVD endpoint was defined as death from any non-cardiovascular cause not included in the primary outcome. Details of the different ICD-10 codes included in both the fatal and non-fatal components of the endpoint are provided in **Supplementary Table 2**.

LIFE-CVD2 was recalibrated to four risk regions of Europe using methods and values previously applied for recalibration of the SCORE2 and SCORE2-OP models, with adaptation for the lifetime risk context (**Supplementary Methods**).[8,9] Briefly, risk regions were defined according to CVD mortality obtained from WHO mortality statistics (**Supplementary Figure 1**) and recalibration involved the use of region-, 5 year age group-, and sex-specific CVD

8

incidence, non-CVD mortality incidence and average risk factors values. To obtain CVD incidence estimates that included non-fatal as well as fatal events, annual CVD mortality rates derived from the WHO mortality database[17] were converted to total CVD rates using a 'multiplier approach', applying multipliers previously published for the SCORE2 and SCORE2-OP risk algorithms.[8,9] For Non-CVD deaths, annual mortality rates derived from the WHO mortality database[8,9] corrected for the presence of individuals with established ASCVD in national mortality statistics using similar multiplier approach. Age group- and sex-specific risk factor values were obtained from the Non-Communicable Disease Risk Factor Collaboration (NCD-RisC).[18,19] Sex-specific recalibration of both the CVD and non-CVD mortality risk models was performed by regressing the observed annual risks on the model predicted 1-year risks (using age group-specific risk factor values) across 5-year age groups (**Supplementary Methods**). The recalibrated CVD and non-CVD death risks were then combined in the lifetable.

Discrimination was assessed using Harrell's C-indices corrected for competing risks.[20] Calibration was assessed by visual inspection of predicted versus observed risk plots in deciles of predicted risk using external validation data from CPRD and ELAN, as these were the only cohorts deemed approximately nationally representative.[20] Since none of the cohorts used for external validation had lifelong follow-up, discrimination and calibration was assessed at 10 years of follow-up, or at the latest full year with at least 80% of the individuals still under follow up, using time in study as the time scale. Calibration of lifetime risk predictions was also assessed in CPRD using age as the timescale.

The handling of missing data is more extensively described in the **Supplementary Methods**. In brief, predictors were imputed by single regression imputation with predictive mean matching for all cohort data. As the routine care data sources (CPRD and ELAN) had much higher rates of missing data, which were more likely to be associated to CVD outcomes, these were handled using multiple imputation with fully conditional specification in 5 imputed

9

datasets. All analyses were performed with R-statistical programming (version 3.5.2, R Foundation for Statistical Computing, Vienna, Austria) or Stata (version 15.1, StataCorp, College Station, Texas, USA). All R and Stata code is available from the corresponding author upon request.

Estimation of preventive intervention effects

The alteration of patient characteristics in individual predictions to simulate treatment effects provides an observational, rather than causal, assessment of risk-factor changes.[21] Therefore, to estimate the lifelong effect of risk factor management (blood pressure lowering, lipid lowering and, smoking cessation), causal evidence was used from trials and meta-analyses, [5,21] which were combined with the annual risks of CVD events and non-CVD mortality as estimated in the updated LIFE-CVD2 model. This process is described in detail in the **Supplementary Methods**. Accordingly, the potential gain in CVD-free life expectancy was estimated by applying causal HRs of 0.78 per 1 mmol/L reduction in LDL cholesterol[22,23] and 0.80 per 10 mm Hg SBP reduction[24] to the CVD event rates. The benefit of smoking cessation was applied using an HR of 0.60 for CVD event rates and 0.73 for non-CVD mortality rates.[25,26] A stable hazard ratio was assumed over time in the examples of the lifetime benefit of risk factor reduction.

Results

Model derivation

Model derivation involved 687,135 individuals without previous CVD recruited between 1990 and 2009 into prospective cohorts in Europe (36 cohorts, 610,353 participants) and North America (9 cohorts, 76,782 participants). Mean age at recruitment was 57 (SD 9) years, 298,408 (43%) were male (**Table 1**). During median follow-up of 10.7 (5th, 95th percentile; 5.0, 18.5) years, a total of 30,939 CVD events and 34,284 non-CVD deaths were recorded. Model HRs are summarised in **Table 2**, with the unrounded log hazard ratios and baseline

hazard shown in **Supplementary Tables 3 and 4**. The HRs for most risk predictors decreased with increasing age of participants.

Using the age-, sex-, and region-specific mean risk factor levels and incidence data, the LIFE-CVD2 model was recalibrated to four European risk regions (**Supplementary Figure 2**). After recalibration, predicted risks based on mean risk factor levels showed good agreement with the expected incidences of CVD event and non-CVD mortality (**Supplementary Figure 3 and 4**) and were also similar to incidence rates obtained from external national registries (**Supplementary Figure 5**). Regional sex- and age-specific multipliers for conversion of CVD mortality rates to incidence rates and for correction of those with prior CVD in national mortality statistics are shown in **Supplementary Figure 6 and 7**. The ratio between the 1-year cumulative incidence of total to fatal CVD was similar to the ratio at 10-years (**Supplementary Figure 8**) supporting use of the SCORE2 project multipliers previously estimated to convert 10-year CVD mortality to total CVD event risk.[8]

External validation

External validation involved data from 1,657,707 individuals without previous CVD or DM from 8 European cohorts. Of these individuals, 793,454 were male (48%) and the mean (SD) ages per cohort ranged from 48 (13) years in the Estonian Biobank to 59 (6) years in HNR (**Supplementary Table 6**). The median follow-up times per cohort ranged from 6.3 years (IQR 6.0-6.9) in HAPIEE Poland to 14.0 years (IQR 10.5-15.6) in HNR. During this follow-up, a total of 61,311 CVD events and 65,867 non-CVD deaths were recorded. The pooled C-index for the prediction of CVD events was 0.795 (95%CI 0.767-0.822) in the studies including the full LIFE-CVD2 age range, the overall pooled estimate including all studies was 0.789 (95%CI 0.703-0.875). (**Figure 2**). The pooled C-index for the competing endpoint of non-CVD mortality was 0.831 (95%CI 0.810-0.852) in the studies with the full age range and 0.824 (95%CI 0.711-0.937) when including all external validation studies. (**Supplementary Figure 9**). C-indices for both endpoints were lower for cohorts which did not include patients over the complete

11

age-range at baseline. In the CPRD and ELAN data, the predicted 10-year CVD event risks agreed well with the observed risks (**Figure 3**), whereas there was some overestimation of non-CVD mortality risk in ELAN (**Supplementary Figure 10**). Sensitivity analysis evaluating the accuracy of lifetime risks in CPRD, showed good agreement between predicted lifetime risk and the observed lifetime incidence of CVD for both men and women (**Supplementary Figure 11**). The LIFE-CVD2 individual predicted risks were well aligned with SCORE2 predictions. For example, among individuals below 70 in CPRD the median difference between the SCORE2 and LIFE-CVD2 10-year predictions was -0.15% (IQR -0.36%; -0.02%).

Estimation of treatment effects

Figure 4 shows how the estimated gain in CVD-free life expectancy, conditional on age, from lifelong 10 mm Hg blood pressure reduction estimated by the LIFE-CVD2 model differs across regions for an individual person with a systolic blood pressure of 140 mm Hg, total cholesterol of 5.5 mmol/L and HDL cholesterol of 1.3 mmol/L. The estimated gain in CVD-free life expectancy in 40-year old, non-smoking men ranged from 0.9 years in low risk countries to 1.7 years in very high-risk countries. In women, it ranged from 0.9 years in low risk countries to 1.6 years in very high-risk countries (**Figure 4**). In comparison, the individual gain in CVD-free life expectancy from smoking cessation in the same 40-year women ranged from 4.2 years in the low risk region up to 4.8 years for 40-year old in the very high risk region, and for 40-year old men from 5.0 in the low risk region up to 6.3 year in the very high risk region (**Supplementary Figure 12**). An example of several 10-year and lifetime prediction measures have been illustrated across different ages in **Supplementary Figure 13**.

The estimated individual gain in CVD-free life expectancy for each combination of risk factor levels is displayed in two-dimensional risk charts in **Appendix 1** showing the examples of 40% mmol/L LDL cholesterol reduction, SBP reduction to <140 mm Hg, and smoking cessation in each of the four European risk regions. The highest gains in CVD-free life expectancy from prevention are in younger individuals with adverse risk factor values, who are often below
12

recommended treatment thresholds. For example, a non-smoking of 52 year woman, with a SBP of 150 mm Hg, total cholesterol of 6.0 mmol/L and HDL cholesterol of 1.1 mmol/L, would have SCORE2 10-year risks of 3.0% in the low risk region ranging up until 8.2% in the very high risk region, all below the guideline-recommended treatment thresholds.[4,8] Based on the LIFE-CVD2 model her gain in gains CVD-free life expectancy from 1 mmol/L LDL reduction would range from 1.1 year in the low risk region until 1.9 years in the very high risk region – a higher benefit than several other individuals above 10-year risk-based treatment thresholds (**Appendix 1)**.[4]

Discussion

This report describes the development of LIFE-CVD2, a risk model to estimate lifetime risk and CVD-free life expectancy in individuals without previous CVD or DM in four European risk regions. It also demonstrates how LIFE-CVD2 can be used to illustrate potential gains in CVD-free life expectancy in response to preventative intervention, given reasonable assumptions about intervention effects. Recalibration was completed using an approach adapted from SCORE2to develop 10-year CVD risk models for the same four risk regions. External validation was performed in across different risk regions and estimations of lifetime treatment benefits were illustrated for several risk factor profiles. The updated LIFE-CVD2 model confers several advantages over the originally published version of the model (LIFE-CVD).

First, it is systematically recalibrated using contemporary and representative data on CVD incidence and risk factor data, which broadens the generalizability of the LIFE-CVD2 model across European risk regions. Because the recalibration approach was based on registry data, the model can be readily updated to reflect future disease CVD incidences and risk factor profiles as soon as new updated data become available.[7,8] Furthermore, since the recalibration approach was aligned with that used in the 10-year CVD SCORE2 and SCORE2-OP risk prediction models, synergy between 10-year and lifetime risk assessment is ensured.

Second, because models have been derived and recalibrated to be sex-specific, LIFE-CVD2 is well-adapted to the contemporary clinical practice for both sexes. The original LIFE-CVD model was not derived and recalibrated separately for both sexes, ignoring differences in the relative effects of certain predictors, and different evolution of risk with age between men and women.

Third, the models have been derived again using powerful, contemporary data from multiple different studies and registry sources. This enhances the accuracy, generalisability and validity of the approach. In particular, data on a total of more than 12.5 million individuals from dozens of countries were used for the development and recalibration of LIFE-CVD2.

Fourth, the age range of the model has been extended from 45-90 to the age range of 35-100 years. This allows the model to be applied to individuals with a current age between 35 and 90 years and improves the stability of estimates in people of all ages with a very high life-expectancy. As the worldwide life expectancy continues to rise,[27] this will be increasingly important.

Fifth, to improve risk communication, the lifetime treatment benefit, defined as the gain in CVD-free life expectancy from preventive therapy, can be estimated with the LIFE-CVD2 model. This has been shown to be an intuitive measure that lowers the decisional conflict among individuals considering preventive treatment.[28] When using lifetime treatment benefit measures in the shared decision process, these should be weighed against the intended treatment duration. All of these of risk estimates should be taken into account when considering treatment initiation, and should be used in conjunction with assessing potential risk modifiers relevant to the specific patient as well as patient preferences.

Finally, the current study highlights the importantly different expected gain in CVD-free life expectancy differed across geographical locations in Europe. This has been incorporated into

14

benefit estimation by the integral recalibration step in model development, something other lifetime risk models have not considered.[21,29]. Our results have shown that interventions are expected to lead to a higher absolute treatment benefit in Eastern European countries, reflecting higher disease incidences in this region.

The LIFE-CVD2 model can be used in a simplified form via the two-dimensional risk charts as provided in Appendix 1, predicting the benefit from lifelong lipid lowering, blood pressure lowering, and smoking cessation. However, to accommodate more accurate predictions and to calculate a wider range of possible treatment options, the LIFE-CVD2 model will be integrated in the CE-marked U-Prevent medical device, available from www.U-Prevent.com. Because of the time required for implementation in a CE-marked medical device, the LIFE-CVD2 models is integrated in an R-shiny app for scientific purposes only (i.e. no clinical use) from https://hagemanshj.shinyapps.io/LIFECVD2/.

The potential limitations of this study merit consideration. Calibration of the LIFE-CVD2 model was only assessed in the large nationally representative dataset from the CPRD and ELAN, because the other cohorts used for external validation do not necessarily reflect contemporary absolute risk levels across European regions (all cohort data, including the cohorts involved may not be nationally representative, reflecting past periods of time or self-selected participants such as healthy volunteers). In the CPRD and ELAN however, good agreement was observed between 10-year predicted and observed CVD incidences. Furthermore, estimated CVD rates agreed well with national incidence rates from available independent external registries from several countries and the discrimination was evaluated in all European risk regions.

Another potential limitation of the current study is that data on medication use, family history, socio-economic status, nutrition, physical activity, renal function, or ethnicity were not available in cohorts and registries used for model derivation and recalibration. Hence,

15

interpretation of LIFE-CVD2 estimates may require clinical judgement, especially for individuals in whom these factors may be relevant.

For the derivation of the LIFE-CVD2 model, data from relatively high-income countries were included; Ideally, however, the derivation of risk models 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 generally exist. Indeed, even in low- and moderate-risk regions, the cohorts involved may not be nationally representative, reflecting past periods of time or self-selected participants such as healthy volunteers. While healthy volunteer bias can lead to low estimates of absolute risk, relative risks are generally unaffected.[7] Furthermore, our approach makes the assumption that the relative risks obtained in the derivation dataset are transferable across different populations, which was further supported by the satisfactory discrimination results in high and very high risk region validation cohorts as observed in the current study.

Another potential limitation is the fact that validation was mostly performed with 10-year risks, as it is not feasible to perform validation of life expectancy measures within the scope of cohort follow-up durations. In CPRD sensitivity analyses using age as the time-scale were performed and showed adequate agreement between predicted and observed lifetime risks. However, similar methodological adaptations to calculate discrimination are not possible when taking a lifetime perspective. Previous studies have shown the validity of CVD risk predictions for up to 17 years.[16] When longer-term data become available, the model could profit from validations at even longer timescales to further validate the underlying methodology.

In conclusion, by taking into account geographical differences in CVD incidence, the recalibrated LIFE-CVD2 model provides a more accurate tool for the prediction of lifetime risk and CVD-free life expectancy for individuals without previous CVD across Europe, ensuring

16

synergy with 10-year risk estimation and facilitating shared decision-making on Step 2 cardiovascular prevention options as recommended by the 2021 European Prevention Guidelines.

17

Data availability

The data underlying this article were provided by representatives of all included cohorts. Data from each cohort may be shared on request to the respective representatives, depending on cohort-specific policies. All R and Stata code is available from the corresponding author upon request.

Author contributions

SHJH, SK, TIV, JAND, EDA, FLJV, LP contributed to the conception or design of the work. SHJH, SK, WL, JMK, BS, LP, KL contributed to the analyses. HJAvO, HP, RK, SM, AP, AT, AS, SSc, TRB, SSp, SJLB, MJB, JMAB, AB, HB, EJB, NRC, KWD, ED, CD, MD, JSF, IF, MF, RTG, SG, RFG, SK, WL, JMK, BS, LP, KL, AGdlC, LLH, POH, PPH, SEH, MKI, JWJ, MK, SKi, AKN, DLP, AMI, KM, HEM, KGMM, MBM, MM, BGN, CP, LPa, APe, DP, LPo, RP, BMP, PMR, BLR, AR, NS, BSc, JES, SS, MJS, RS, BT, WMMV, HV, NJW, LW, PW, BZ, JD, MEN, MB, RE, contributed to the data acquisition of at least one of the included cohorts, SHJH drafted the manuscript, which was critically revised by all co-authors. All gave final approval and agree to be accountable for all aspects of the work ensuring integrity and accuracy.

References

1. Roth GA, Abate D, Abate KH, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*. 2018;392(10159):1736-1788. doi:10.1016/S0140-6736(18)32203-7

2. van der Leeuw J, Ridker PM, van der Graaf Y, Visseren FLJ. Personalized cardiovascular disease prevention by applying individualized prediction of treatment effects. *Eur Heart J*. 2014;35(13):837-843. doi:10.1093/eurheartj/ehu004

3. Dorresteijn JAN, Visseren FLJ, Ridker PM, et al. Estimating treatment effects for individual patients based on the results of randomised clinical trials. *BMJ*. 2011;343(oct03 1):d5888-d5888. doi:10.1136/bmj.d5888

4. Visseren FLJ, Mach F, Smulders YM, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. *Eur Heart J*. 2021;42(34):3227-3337. doi:10.1093/eurheartj/ehab484

5. Jaspers NEM, Blaha MJ, Matsushita K, et al. Prediction of individualized lifetime benefit from cholesterol lowering, blood pressure lowering, antithrombotic therapy, and smoking cessation in apparently healthy people. *Eur Heart J*. 2019;31:1-10. doi:10.1093/eurheartj/ehz239

6. Pennells L, Kaptoge S, Wood A, et al. Equalization of four cardiovascular risk algorithms after systematic recalibration: individual-participant meta-analysis of 86 prospective studies. *Eur Heart J*. 2019;40(7):621-631. doi:10.1093/eurheartj/ehy653

7. Kaptoge S, Pennells L, De Bacquer D, et al. World Health Organization cardiovascular disease risk charts: revised models to estimate risk in 21 global regions. *Lancet Glob Health*. 2019;7(10):e1332-e1345. doi:10.1016/S2214-109X(19)30318-3

8. Hageman S, Pennells L, Ojeda F, et al. SCORE2 risk prediction algorithms: new models to estimate 10-year risk of cardiovascular disease in Europe. *Eur Heart J*. 2021;42(25):2439-2454. doi:10.1093/eurheartj/ehab309

9. de Vries TI, Cooney MT, Selmer RM, 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(25):2455-2467. doi:10.1093/eurheartj/ehab312

10. Danesh J, Erqou S, Walker M, 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(12):839-869. doi:10.1007/s10654-007-9165-7

11. Hewitt J, Walters M, Padmanabhan S, Dawson J. Cohort profile of the UK Biobank: diagnosis and characteristics of cerebrovascular disease. *BMJ Open*. 2016;6(3):e009161-e009161. doi:10.1136/bmjopen-2015-009161

12. Kist JM, Vos RC, Mairuhu ATA, et al. SCORE2 cardiovascular risk prediction models in an ethnic and socioeconomic diverse population in the Netherlands: an external validation study. *eClinicalMedicine*. 2023;57:101862. doi:10.1016/j.eclinm.2023.101862

13. Schmermund A, Möhlenkamp S, Stang A, et al. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: Rationale and design of the Heinz Nixdorf RECALL study. *Am Heart J.* 2002;144(2):212-218. doi:10.1067/mhj.2002.123579

14. Leitsalu L, Haller T, Esko T, et al. Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. *Int J Epidemiol.* 2015;44(4):1137-1147. doi:10.1093/ije/dyt268

15. Peasey A, Bobak M, Kubinova R, et al. Determinants of cardiovascular disease and other non-communicable diseases in Central and Eastern Europe: Rationale and design of the HAPIEE study. *BMC Public Health.* 2006;6(1):255. doi:10.1186/1471-2458-6-255

16. Dorresteijn JAN, Kaasenbrood L, Cook NR, et al. How to translate clinical trial results into gain in healthy life expectancy for individual patients. *BMJ.* 2016;352:i1548. doi:10.1136/bmj.i1548

17. World Health Organization. WHO Mortality Database. Accessed May 7, 2020. https://apps.who.int/healthinfo/statistics/mortality/whodpms/

18. NCD Risk Factor Collaboration. Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19·1 million participants. *The Lancet.* 2017;389(10064):37-55. doi:10.1016/S0140-6736(16)31919-5

19. NCD Risk Factor Collaboration. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4·4 million participants. *The Lancet.* 2016;387(10027):1513-1530. doi:10.1016/S0140-6736(16)00618-8

20. Wolbers M, Koller MT, Witteman JCM, Steyerberg EW. Prognostic Models With Competing Risks. *Epidemiology.* 2009;20(4):555-561. doi:10.1097/EDE.0b013e3181a39056

21. Lloyd-Jones DM, Huffman MD, Karmali KN, et al. Estimating Longitudinal Risks and Benefits From Cardiovascular Preventive Therapies Among Medicare Patients. *J Am Coll Cardiol.* 2017;69(12):1617-1636. doi:10.1016/j.jacc.2016.10.018

22. Sabatine MS, Giugliano RP, Keech AC, et al. Evolocumab and Clinical Outcomes in Patients with Cardiovascular Disease. *N Engl J Med.* 2017;376(18):1713-1722. doi:10.1056/NEJMoa1615664

23. Baigent C, Blackwell L, Emberson J, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: A meta-analysis of data from 170 000 participants in 26 randomised trials. *The Lancet.* 2010;376(9753):1670-1681. doi:10.1016/S0140-6736(10)61350-5

24. Ettehad D, Emdin CA, Kiran A, et al. Blood pressure lowering for prevention of cardiovascular disease and death: A systematic review and meta-analysis. *The Lancet.* 2016;387(10022):957-967. doi:10.1016/S0140-6736(15)01225-8

25. Mons U, Muezzinler A, Gellert C, et al. Impact of smoking and smoking cessation on cardiovascular events and mortality among older adults: meta-analysis of individual participant data from prospective cohort studies of the CHANCES consortium. *BMJ.* 2015;350(apr20 2):h1551-h1551. doi:10.1136/bmj.h1551

26. Gellert C, Schöttker B, Brenner H. Smoking and All-Cause Mortality in Older People. *Arch Intern Med.* 2012;172(11):837-844. doi:10.1001/archinternmed.2012.1397

27. World Health Organization. Life expectancy and Healthy life expectancy: Data by WHO region. Accessed December 3, 2021. https://www.who.int/data/gho/data/themes/topics/indicator-groups/indicator-group-details/GHO/life-expectancy-and-healthy-life-expectancy

28. Jaspers NEM, Visseren FLJ, van der Graaf Y, et al. Communicating personalised statin therapy-effects as 10-year CVD-risk or CVD-free life-expectancy: does it improve decisional conflict? Three-armed, blinded, randomised controlled trial. *BMJ Open.* 2021;11(7):e041673. doi:10.1136/bmjopen-2020-041673

29. Hippisley-Cox J, Coupland C, Robson J, Brindle P. Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database. *BMJ.* 2010;341(dec09 1):c6624-c6624. doi:10.1136/bmj.c6624

Table 1: Summary statistics of the model derivation population

|	N (%) or mean (SD)
Total participants	687,135
Male sex	298,408 (43%)
Age (years)	58 (8)
Current smoker	104,471 (15%)
Systolic blood pressure (mmHg)	135 (19)
Diabetes mellitus	32,234 (4.7%)
Total cholesterol (mmol/l)	5.9 (1.1)
HDL-cholesterol (mmol/l)	1.4 (0.4)
Follow-up (years, 5th/95th percentile)	10.7 (5.0-18.5)
Cardiovascular events	30,939
Non-cardiovascular deaths	34,284

HDL = high density lipoprotein

Table 2: Hazard ratios of the LIFE-CVD2 models

	Men		Women	
Models to predict CVD events	**Main effect**	**Age interaction term (per 5 year)**	**Main effect**	**Age interaction term (per 5 year)**
Age (per 5 years)	1.10 (1.08, 1.13)		1.13 (1.10, 1.16)	
Current smoking (versus never/former)	1.90 (1.83, 1.97)	0.94 (0.92, 0.96)	2.27 (2.17, 2.38)	0.90 (0.88, 0.93)
SBP (per 20mmHg)	1.34 (1.31, 1.36)	0.96 (0.96, 0.97)	1.40 (1.37, 1.43)	0.95 (0.94, 0.96)
Total cholesterol (per 1 mmol/L)	1.16 (1.15, 1.18)	0.97 (0.96, 0.98)	1.11 (1.09, 1.13)	0.97 (0.96, 0.98)
HDL cholesterol (per 0.5 mmol/L)	0.76 (0.74, 0.78)	1.04 (1.03, 1.06)	0.78 (0.76, 0.80)	1.04 (1.03, 1.05)
History of diabetes mellitus*	1.94 (1.84, 2.04)	0.91 (0.89, 0.94)	2.35 (2.21, 2.50)	0.89 (0.86, 0.91)
Models to predict non-CVD mortality	**Main effect**	**Age interaction term (per 5 year)**	**Main effect**	**Age interaction term (per 5 year)**
Age (per 5 years)	0.99 (0.97, 1.02)		0.90 (0.88, 0.92)	
Current smoking (versus never/former)	2.22 (2.14, 2.30)	0.93 (0.92, 0.95)	2.18 (2.10, 2.27)	0.98 (0.97, 1.00)
SBP (per 20mmHg)	1.08 (1.06, 1.10)	0.96 (0.96, 0.97)	1.06 (1.04, 1.08)	0.98 (0.97, 0.99)
Total cholesterol (per 1 mmol/L)	0.92 (0.90, 0.93)	1.00 (0.99, 1.01)	0.95 (0.94, 0.97)	0.98 (0.98, 0.99)
HDL cholesterol (per 0.5 mmol/L)	1.10 (1.08, 1.13)	0.98 (0.97, 0.99)	0.95 (0.93, 0.97)	1.01 (1.00, 1.03)
History of diabetes mellitus*	1.61 (1.52, 1.70)	0.92 (0.89, 0.94)	1.72 (1.61, 1.83)	0.94 (0.91, 0.97)

Sex-specific hazard ratios from the LIFE-CVD2 models predicting the risk of fatal and non-fatal CVD events or the risk of non-CVD mortality. (Baseline) age was centered at 60 years, systolic blood pressure at 120 mmHg, total cholesterol at 6 mmol/L, and HDL cholesterol at 1.3 mmol/L. HDL = high density lipoprotein, SBP = systolic blood pressure.

*Diabetes mellitus was included in the modelling as diabetes patients are included in the recalibration data. For use in clinical practice this coefficient should be ignored.
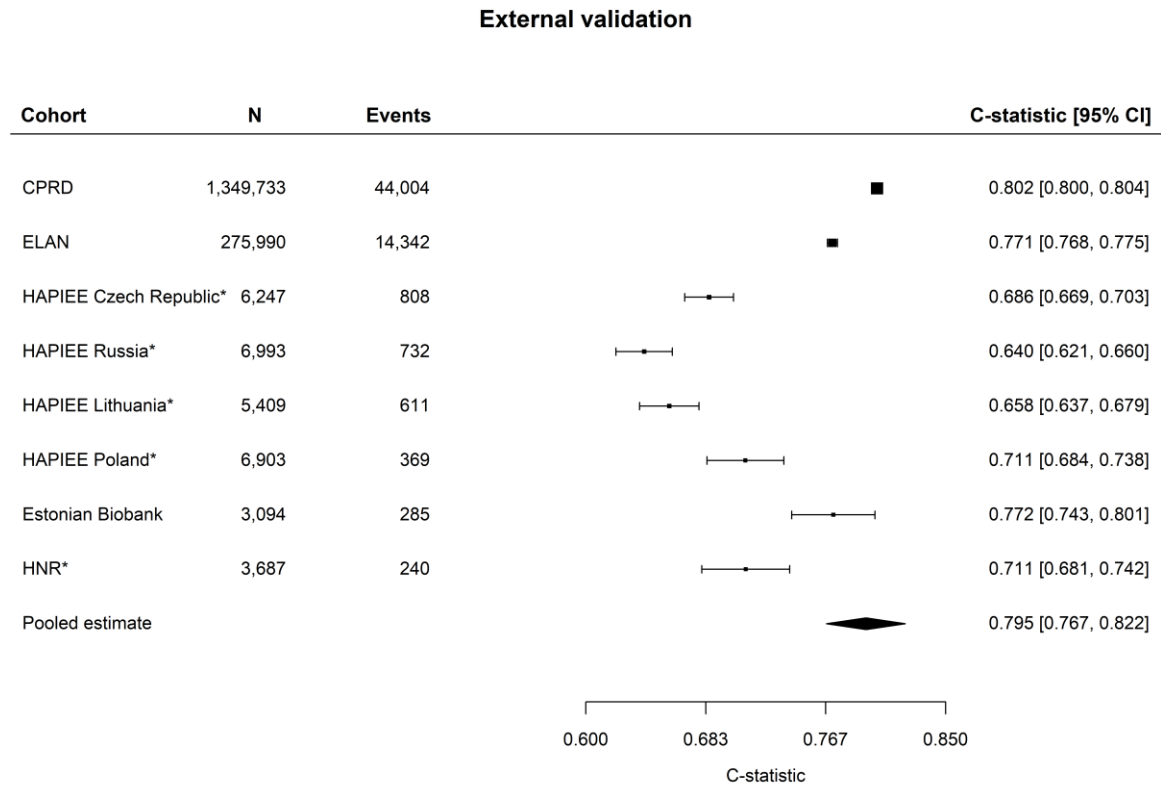
Figure 1: Study design
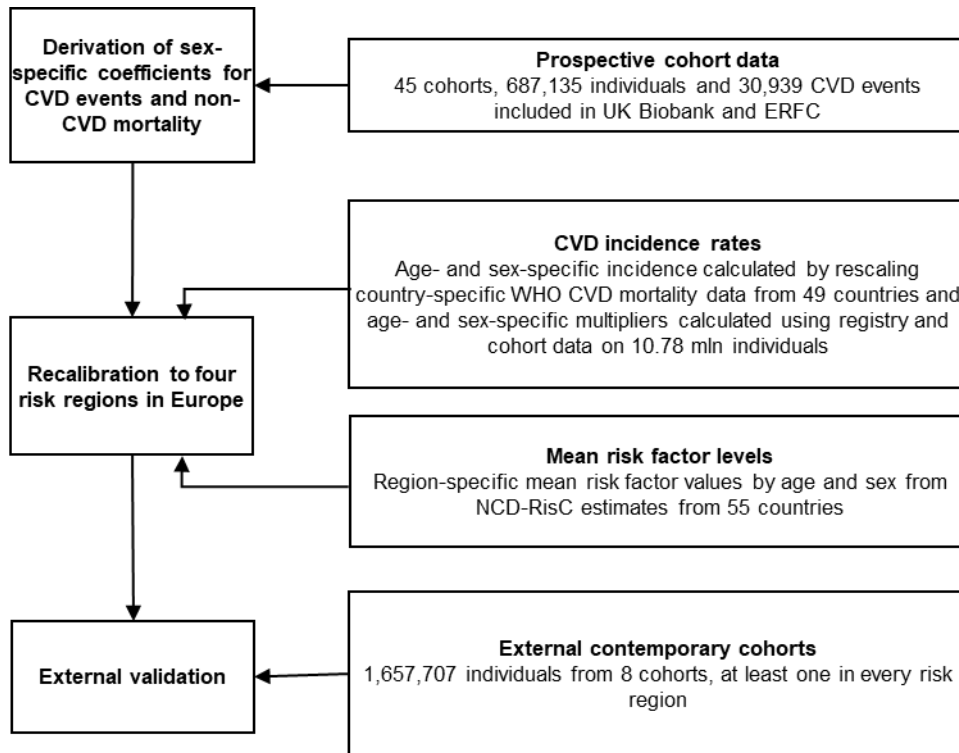
CVD = cardiovascular disease, ERFC = Emerging Risk Factor Collaboration, WHO = World Health Organisation, NCD RisC = NCD Risk Factor Collaboration
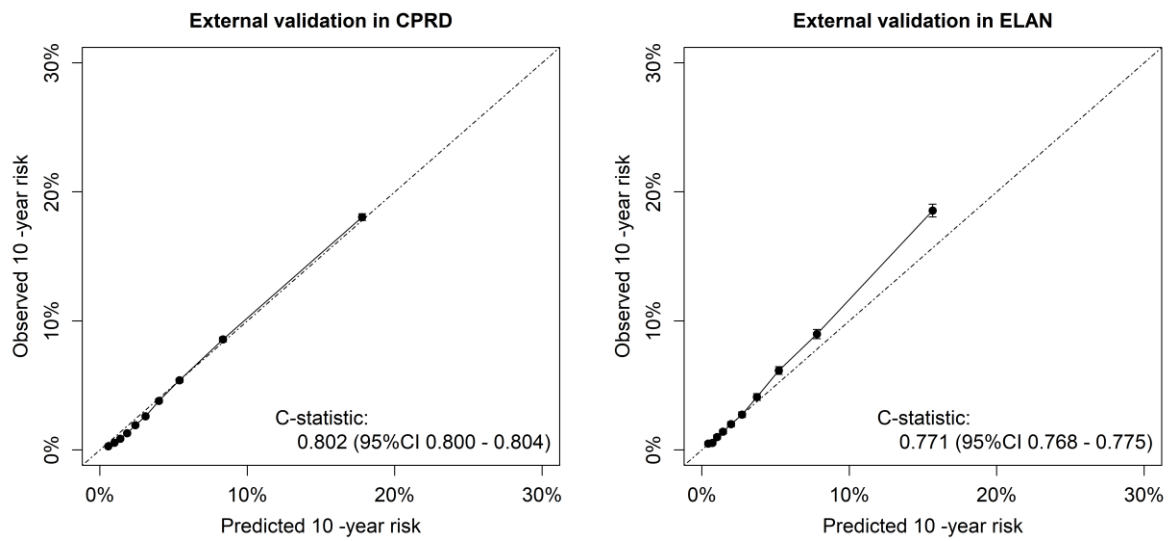
Figure 2: C-index of the recalibrated LIFE-CVD model to discriminate in external validation cohorts upon assessing CVD events

External validation

| Cohort	N	Events		C-statistic [95% CI]
CPRD	1,349,733	44,004		0.802 [0.800, 0.804]
ELAN	275,990	14,342		0.771 [0.768, 0.775]
HAPIEE Czech Republic*	6,247	808		0.686 [0.669, 0.703]
HAPIEE Russia*	6,993	732		0.640 [0.621, 0.660]
HAPIEE Lithuania*	5,409	611		0.658 [0.637, 0.679]
HAPIEE Poland*	6,903	369		0.711 [0.684, 0.738]
Estonian Biobank	3,094	285		0.772 [0.743, 0.801]
HNR*	3,687	240		0.711 [0.681, 0.742]
Pooled estimate				0.795 [0.767, 0.822]

C-statistic axis: 0.600   0.683   0.767   0.850

* = as discrimination result may be underestimated because the entire age range of the LIFE-CVD2 model could not be included from this cohort, this result is not included in the shown pooled estimate. The overall pooled estimate including all studies is 0.789 (95%CI 0.703-0.875)

Figure 3: Calibration for prediction CVD events of the recalibrated LIFE-CVD2 model in CPRD (n=1,349,377) and ELAN (n=275,990)

Predicted versus observed CVD risks in deciles of predicted risk for the recalibrated LIFE-CVD2 model.

Figure 4: Predicted gain in CVD-free life expectancy from 10 mm Hg blood pressure reduction for an individual with total cholesterol concentrations of 5.5 mmol/L, HDL cholesterol of 1.3 mmol/L, and systolic blood pressure of 140 mm Hg, for each region, stratified on smoking status.
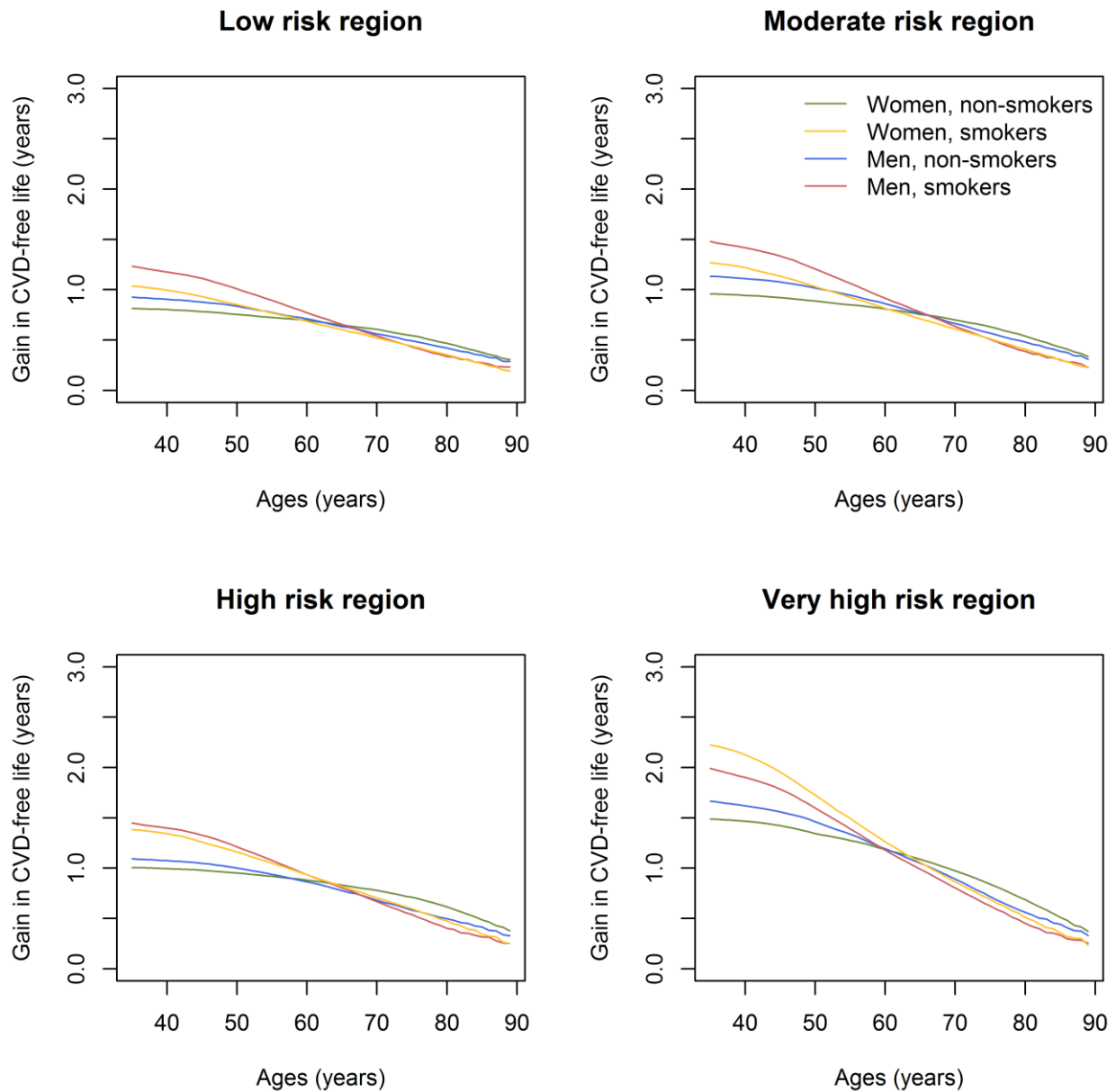
Author list

Steven HJ Hageman* (Department of Vascular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands), Stephen Kaptoge* (British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge UK; Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge UK), Tamar I de Vries (Department of Vascular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands), Wentian Lu (Department of Epidemiology and Public Health, University College London, London, UK), Janet M Kist (Health Campus The Hague and Department of Public Health & Primary Care, Leiden University Medical Center, the Netherlands), Kristi Läll (Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia), Börge Schmidt (Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University Duisburg-Essen, Essen, Germany), Hendrikus JA van Os (Department of Public Health & Primary Care and National eHealth Living Lab, Leiden University Medical Center, the Netherlands), Hynek Pikhart (Department of Epidemiology and Public Health, University College London, London, UK, RECETOX, Masaryk University, Czech Republic)
Ruzena Kubinova (National Institute of Public Health, Czech Republic), Sofia Malyutina (Independent researcher), Andrzej Pająk (Department of Epidemiology and Population Studies, Institute of Public Health, Faculty of Health Sciences, Jagiellonian University Medical College, Kraków, Poland), Abdonas Tamosiunas (Institute of Cardiology, Lithuanian University of Health Sciences, Kaunas, Lithuania), Andreas Stang (Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University Duisburg-Essen, Essen, Germany), Sara Schramm (Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University Duisburg-Essen, Essen, Germany), Thomas R Bolton (British Heart Foundation Data Science Centre, Health Data Research UK, London, UK; British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK), Sarah Spackman (British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge UK; Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge UK), Stephan J.L. Bakker (Department of Internal Medicine, University Medical Centre Groningen, University of Groningen, Groningen, Netherlands), Michael J Blaha (Johns Hopkins Ciccarone Center for the Prevention of Heart Disease, Johns Hopkins Hospital, Baltimore, MD, USA), Jolanda M A Boer (Centre for Prevention, Lifestyle and Health, National Institute for Public Health and the Environment, Bilthoven, The Netherlands), Amélie Bonnefond (Inserm/CNRS UMR 1283/8199, Pasteur Institute of Lille, EGID, Lille, France; University of Lille, Lille, France; Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK), Hermann Brenner (Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany. Network Aging Research, University of Heidelberg, Heidelberg, Germany), Eric J Brunner (Department of Epidemiology and Public Health University College London, London, UK), Nancy R Cook (Brigham & Women's Hospital, Harvard Medical School Harvard University, Boston, MA, USA), Karina W Davidson (Feinstein Institutes for Medical Research, Northwell Health, New York, NY, USA), Elaine Dennison (MRC Lifecourse Epidemiology Centre, University of Southampton, Southampton, UK), Chiara Donfrancesco (Department of Cardiovascular, Endocrine-Metabolic Diseases and Aging, Istituto Superiore di Sanita', Rome, Italy), Marcus Dörr (Institute for Community Medicine, University Medicine Greifswald, University of Greifswald, Greifswald, Germany. German Centre for Cardiovascular Disease (DZHK), Partner Site Greifswald, and German Centre for Cardiovascular Disease (DZD), Site Greifswald, Greifswald, Germany) James S Floyd (Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle, WA USA), Ian Ford (Robertson Center for Biostatistics, University of Glasgow, United Kingdom), Michael Fu (Dept of Medicine, Sahlgrenska University Hospital/Östra Hospital, Gothenburg, Sweden), Ron T. Gansevoort (Department of Internal Medicine, University Medical Centre Groningen, University of Groningen, Groningen, Netherlands), Simona Giampaoli (former Istituto Superiore di Sanità, Rome, Italy), Richard F Gillum (Howard University Hospital, Washington DC, DC, USA), Agustín Gómez-de-la-Cámara (Instituto de Investigación Hospital 12 de Octubre, Madrid, Spain), Lise Lund Håheim (Institute of Oral Biology, Faculty of Dentistry, University of Oslo, Norway), Per-Olof Hansson (Department of Molecular and Clinical Medicine, Institute of Medicine, University of Gothenburg, Sahlgrenska Academy, Gothenburg, Sweden), Peter P Harms (Department of General Practice, Amsterdam University Medical Center, Amsterdam, Netherlands), Steve E Humphries (Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, United Kingdom), M Kamran Ikram (Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, Netherlands), J. Wouter Jukema (Department of Cardiology, Leiden University Medical Center, The Netherlands, Netherlands Heart Institute, Leiden, the Netherlands), Maryam Kavousi (Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, Netherlands), Stefan Kiechl (Department of Neurology, Innsbruck Medical University and VASCage, Research Centre on Vascular Ageing

and Stroke, Innsbruck, Austria), Anna Kucharska-Newton (College of Public Health, Department of Epidemiology, University of Kentucky, KY, USA), David Lora Pablos (Instituto de Investigación Hospital 12 de Octubre, Universidad Complutense de Madrid (UCM), Spain), Alejandro Marín Ibañez (San Jose Norte Health Centre, Zaragoza, Spain), Kunihiro Matsushita (Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA), Haakon E Meyer (Norwegian Institute of Public Health, Oslo, Norway), Karel G M Moons (Julius Center for Health Sciences and Primary Care, UMC Utrecht, Utrecht University, Utrecht, The Netherlands), Martin Bødtker Mortensen (Department of Cardiology, Aarhus University Hospital, Denmark), Mirthe Muilwijk (Department of Epidemiology and Data Science, Amsterdam University Medical Center, Amsterdam, Netherlands), Børge G Nordestgaard (Department of Clinical Biochemistry, Copenhagen University Hospital – Herlev Gentofte, Denmark; Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark), Chris Packard (School of Cardiovascular & Metabolic Health, University of Glasgow, United Kingdom), Luigi Pamieri (Department of Cardiovascular, Dysmetabolic and Aging-associated Diseases, Istituto Superiore di Sanità, Rome, Italy), Annette Peters (Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; IBE, Pettenkofer School of Public Health, Medical Faculty, Ludwig-Maximilians-Universität, Munich, Germany; German Centre for Cardiovascular Research (DZHK e.V.), partner site Munich Heart Alliance, Munich, Germany)Demosthenes Panagiotakos (Harokopio University, Athens, Greece), Louis Potier (Université Paris City, Paris, France; Department of Diabetology, Endocrinology and Nutrition, Assistance Publique - Hôpitaux de Paris, Bichat Hospital, Paris, France), Rui Providencia (Institute, of Health Informatics Research, University College London, London, UK), Bruce M Psaty (Cardiovascular Health Research Unit, University of Washington, Seattle, WA, USA), Paul M Ridker (Brigham & Women's Hospital, Harvard Medical School Harvard University, Boston, MA, USA), Beatriz L Rodriguez (University of Hawaii and Tecnologico de Monterrey, Honolulu, HI, USA), Annika Rosengren (Sahlgrenska University Hospital and Östra Hospital, Göteborg, Sweden), Naveed Sattar (BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, UK), Ben Schöttker (Division of Clinical Epidemiology and Aging Research, German Cancer Research Center, Heidelberg, Germany. Network Aging Research, University of Heidelberg, Germany), Joseph E Schwartz (Columbia University, New York, NY, USA), Steven Shea (College of Physicians & Surgeons and Mailman School of Public Health, Columbia University, NY, USA), Martin J Shipley (Department of Epidemiology and Public Health University College London, London, UK), Reecha Sofat (Department of Pharmacology and Therapeutics, University of Liverpool), Barbara Thorand (Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany; IBE, Pettenkofer School of Public Health, Medical Faculty, Ludwig-Maximilians-Universität, Munich, Germany), W M Monique Verschuren (Centre for Prevention, Lifestyle and Health, National Institute for Public Health and the Environment, Bilthoven, The Netherlands and Julius Center for Health Sciences and Primary Care, UMC Utrecht, Utrecht University, Utrecht, The Netherlands), Henry Völzke (Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany. German Centre for Cardiovascular Disease (DZHK), Partner Site Greifswald, and German Centre for Cardiovascular Disease (DZD), Site Greifswald, Greifswald, Germany), Nicholas J Wareham (MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, University of Cambridge, Cambridge, UK), Leo Westbury (MRC Lifecourse Epidemiology Centre, University of Southampton, Southampton, UK), Peter Willeit (Department of Medical Statistics, Informatics and Health Economics, Medical University of Innsbruck, Austria; British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK), Bin Zhou (Faculty of Medicine, School of Public Health, Imperial College London, London, UK), John Danesh (British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge UK; Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge UK), Mattijs E Numans (Health Campus The Hague, Leiden University Medical Center, the Netherlands), Martin Bobak (Department of Epidemiology and Public Health, University College London, London, UK, RECETOX, Masaryk University, Czech Republic), Raimund Erbel (Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, University Duisburg-Essen, Essen, Germany), Frank LJ Visseren† (Department of Vascular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands), Emanuele Di Angelantonio† (British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge UK; Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge, Cambridge UK; British Heart Foundation Centre of Research Excellence, University of Cambridge, Cambridge, UK; National Institute for Health and Care Research Blood and Transplant Research Unit in Donor Health and Behaviour, University of Cambridge, Cambridge, UK; Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge, UK; Health Data Science Research Centre, Human Technopole, Milan, Italy 20157), Lisa Pennells† (British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge UK; Victor Phillip Dahdaleh Heart and Lung

Research Institute, University of Cambridge, Cambridge UK), Jannick AN Dorresteijn† (Department of Vascular Medicine, University Medical Center Utrecht, Utrecht, The Netherlands)

*Contributed equally
† Contributed equally

Investigators

Investigators of the Emerging Risk Factors Collaboration
ATENA: Salvatore Panico, Giovanni Veronesi, Chiara Donfrancesco; **ATTICA**: Christina Chrysohoou; Christos Pitsavos **ARIC:** Anna Kucharska-Newton, Kunihiro Matsushita; **BRUN**: Gregorio Rungger, Christoph Leitner, Johann Willeit; **BWHHS**: Julie Taylor, Caroline Dale, Floriaan Schmidt; **CHS1**: Bruce M Psaty, James S Floyd; **CHS2**: Bruce M Psaty, James S Floyd; **COPEN**: Marianne Benn, Anne Langsted, Pia R Kamstrup; **DESIR**: Beverley Balkau, Fabrice Bonnet, Frederic Fumeron; **DRECE**: Pilar Cancelas-Navia, Cristina Martín-Arriscado, Miguel Menéndez-Orenga, **EMOFRI**: Salvatore Panico, Giovanni Veronesi, Chiara Donfrancesco; **EPICNOR**: Nicholas J Wareham; **ESTHER**: Bernd Holleczek, Kira Trares, Hannah Stocker; **FINNMARK**: Inger Ariansen; **GOTO43**: George Lappas; **HCS**: Cyrus Cooper, Nicholas Harvey; **HONOL**: Beatriz Rodriguez; **HOORN**: Femke Rutters, Marieke T Blom; **HUBRO**: Inger Ariansen; **LEADER**: Jackie A Cooper; **MATISS93**: Benedetta Marcozzi, Cecilia Damiano, Cinzia Lo Noce; **MESA**: Karol Watson, Wendy Post, James Pankow; **MIDFAM**: Naveed Sattar; **MONFRI94**: Salvatore Panico, Giovanni Veronesi, Chiara Donfrancesco; **MOINICA_KORA**: Annette Peters, Barbara Thorand; **MORGEN**: Anneke Blokstra; **MOSWEGOT**: George Lappas; **NHANESIII**: Richard F Gillum; **NSHS**: Karina W Davidson, Joseph E Schwartz; **OPPHED**: Inger Ariansen; **OSLO2**: Lise Lund Håheim; **PREVEND**: Harold Snieder, Kevin Damman; **ProspectEPIC**: Karel G M Moons; **PROSPER**: Stella Trompet, Chris Packard, Ian Ford; **RS_I**: M Arfan Ikram, Banafsheh Arshi, Elif Aribas; **RS_II**: M Arfan Ikram, Banafsheh Arshi, Elif Aribas; **SHIP**: Henry Völzke, Marcello P Markus, Carsten O Schmidt, Stephan B Felix; **TROMS**: Inger Ariansen; **USPHS2**: Julie Buring; **WHITEII**: Eric J Brunner, Mika Kivimaki; **WHITEI**: Eric J Brunner, Mika Kivimaki; **WHS**: Julie Buring; **WOSCOPS**: Michele Robertson, Robin Young, Peter Macfarlane; **ZARAGOZA**: Alejandro Marín Ibañez.