Title: Exploring the strength of relation of cardiovascular health with future cardiovascular events using Life Simple Seven and Life Essential Eight approach

Master of Science degree in Epidemiology (Medicine faculty-Utrecht University): major research project

Name student: Aleksandra Danailova (2024)

Name tutor: prof. Dr. Michiel L. Bots, MD, PhD, University Medical Center Utrecht, Utrecht University

Second reviewer: Dr. Ewoud Schuit, PhD, University Medical Center Utrecht

Project leader / project advisor: prof. Dr. Michiel L. Bots, MD, PhD, Utrecht University

Laymen summary

Cardiovascular diseases (CVDs) are a group of atherosclerosis-related disorders of the heart and blood vessels and are the leading cause of death globally. Millions of people die from CVDs due to heart attacks and strokes. So, tackling these health problems on time is crucial for proper prevention.

The American Heart Association (AHA) developed the so-called Life Simple Seven (LS7) and Life Essential Eight (LE8) measurements to assess the cardiovascular health status of an individual and a population. These metrics are based on health behaviours such as smoking, physical activity [PA], diet, weight and health factors such as cholesterol, blood pressure [BP], glucose control and recently added sleep as an additional factor that contributes to cardiovascular health (CVH). If mitigated via lifestyle changes or interventions, all these factors and health behaviours can prevent CVD. Therefore, LS7 and LE8 serve as crucial instruments for managing the risks of cardiovascular disease (CVD) and gaining insights into promoting healthy ageing. However, there is no in-depth research to determine if the eight (8) components/metrics of CVH (LE8) have the same impact on the risk of developing a CVD event or whether they have different magnitudes and directions on developing atherosclerosis and heart failure as LS7 components. Therefore, studying and presenting the length to which extent these two models differ in predicting the hazard ratio remains a challenge that needs attention.

To assess to which extent the LS7 or LE8 model better relates to the risk of developing CVD, we have quantified cardiovascular health into an ideal, intermediate and poor, performed several analyses on the hazard ratio using the Cox regression model and NRI- Net Reclassification index to compare the two models. We have conducted comprehensive analyses of the hazard ratios at both the total/aggregate level (considering all metrics simultaneously) and the individual level (examining the effect size of each metric, factor or predictor separately) for both models. We performed stratification based on sex and age for cohorts to check the differences in risk prediction based on the two models.

In this study, six (6) out of eight (8) components of the CVH approach were used using almost half of the data from 62.769 in the USE-IMT data set, for which all data regarding health behaviours and factors were available per participant. The results show that the health behaviour indicators and health factors on the total level based on LE8 metrics show a different magnitude of development of CVD compared to the indicators based on LS7, while their direction is the same. Individually, there is a difference in the magnitude and direction of some groups and metrics. Based on the NRI comparison, the new LE8 model is more accurate than the LS7 and the preferred model for use. Further research could build on the findings by investigating the factors contributing to its superior predictive accuracy.

ABSTRACT[1]

CVD severely impacts the quality of Life and mortality (McKearnan, Wolfson, Vock, Vazquez-Benitez & O'Connor, 2018). This calls for urgent preventive measures, so diverse models are proposed to tackle this problem. AHA developed criteria for ideal health dubbed the Life Simple Seven (LS7) and Life Essential Eight (LE8) to quantify population health based on health behaviours and factors (Lloyd-Jones et al., 2022; Lloyd-Jones et al., 2022a). When these factors are optimized, they are associated with better CVD-free survival, total longevity, and improved quality of Life. The project aims to assess to which extent the LS7 or LE8 model better relates to the risk of developing CVD. Using information from the individual participant data of various international population cohorts in the USE-IMT initiative[2], we quantified cardiovascular health into ideal, intermediate/moderate and poor health based on LS7 and LE8 approaches. Cox semi-parametric models were used to estimate the HRr-hazard rates for CVD. The HR-hazard ratios of CVD for the ideal versus the poor category and intermediate versus poor were then calculated separately based on a couple of inputs for LS7 and LE8. The two main models were compared using the net reclassification improvement (NRI) method to explore which model best predicts the risk. We used six (6) components (health behaviours and factors) of the CVH approach using data from 31,549 participants from 16 cohorts (mean age 57.80 yrs, 53% women). LS7's hazard ratio of ideal versus poor health for developing a cardiovascular event is 0.25 with a CI (0.207-0.306), and for the intermediate category versus poor is 0.437 with a CI (0.360-0.530). The corresponding hazard ratio based on LE8 is 0.237 with a CI of (0.209-0.269) and o.468 with a CI of (0.416-0.527) respectively. Individually, the metrics differ in the magnitude and some in the direction of effects. Based on the analysis performed with NRI, the net reclassification improvement index (NRI) is 0.098.

In conclusion: the new LE8 model is slightly more accurate than the LS7, meaning LE8 classified subjects approximately 10% more accurately/correctly than the old model, FS7. So, LE8 is the preferred model for use. Further research could build on the findings by exploring the potential applications of the LE8 model in clinical settings.

INTRODUCTION

According to the World Health Organization (WHO, 2021), cardiovascular diseases are the leading cause of death globally, representing 32% of all global deaths. In 2020, approximately 19.1 million individuals lost their lives due to cardiovascular diseases (CVDs). (Coronado, Melvin, Bell & Zhao, 2022).

In the Netherlands, based on (Koop, Wimmers, Bots and Vaartjes, 2023), ischemic heart disease contributed to total cardiovascular disease mortality by 21% in 2022. Also, there were

Abbriviations: AHA: American Heart Association, BM: Body Mass, BMI:body mass index;, BP: Blood Pressure, CVDs: Cardiovascular Diseases, CVH: Cardiovascular Health, CI: Confidence Interval, dbp: diastolic blood pressure, FBG: fasting blood glucose, HDL: high-density lipoprotein, HbA1c: haemoglobin A1c , HRr: Hazard Rate, HR: Hazard Ratio, LS7: Life Simple Seven, LE8: Life Essential Eight, NDS: nicotine-delivery system, NRI: Net Reclassification Improvement PA: Physical Activity, sbp: systolic blood pressure, BP, blood pressure; SD: Standard deviation, SMQ: smoking assessment.

[2] *USE Intima-Media Thickness is a global meta-analysis project using individual participant data from prospective cohort studies in asymptomatic individuals at risk for cardiovascular disease. Retrieved from [USE-IMT, 2012].*

56,655 hospital admissions, with an average age mean of 67 for men and 71 with ischemic heart disease in 2022. As per other articles and data since 2021, coronary heart disease (CHD) accounts for around 85,000 cases and 23,000 strokes yearly in the Netherlands (Amgen (Europe) GmbH, 2021). According to the same report, an estimated 17,000 people in the Netherlands died from heart disease and 12,000 from stroke in 2020. In 2020, CHD in North Macedonia reached 4,420 or 19% of total deaths, ranking the country number 62 in the world out of 183 according to WHO data since 2020 (WHO, 2021).

CVDs are a group of atherosclerosis-related disorders of the heart and blood vessels. They can include, besides coronary heart disease (CHD), other conditions such as heart failure, cerebrovascular disease, myocardial infarction, peripheral arterial disease, aortic disease, and arrhythmia (Lopez, Ballard, & Jan, 2023). According to Pahwa and Jialal (2023), in the electronic book by Padda, Fabian, & Johal, (2023), atherosclerosis is a chronic arterial condition. The major clinical manifestations of atherosclerosis include coronary heart disease (CHD), ischemic stroke and peripheral arterial disease.

Behaviour and health factors connected to CVD

Modifiable risk factors for atherosclerosis are increased body mass index (BMI), elevated systolic (SBP) or diastolic blood pressure (DBP), hypertension, glucose intolerance, smoking, increased alcohol consumption, low physical activity, dyslipidaemia, type 2 diabetes (treated or untreated) (Herrington et al., 2016). Therefore, beneficial changes in these risk factors on a population and individual level will likely be followed by changes in vascular disease incidence and less CVD (Koopman et al., 2016). Even in 2010, AHA recognized these factors as the main reasons connected to CVH. In 2010, the American Heart Association promoted a positive shift towards health promotion and preservation across the life course in populations and individuals instead of disease treatment (Lloyd-Jones et al., 2022). In 2018, based on the same study, an important update was completed by AHA to improve cardiovascular health (CVH) further in the general population. They included sleep quality (assessed by sleep duration) as an extra factor, in addition to the update on the scoring algorithm for the seven behaviours and health factors (Lloyd-Jones et al., 2022a).

Why Life's Simple 7 [LS7], is relevant for society.

Tsao et al. (2022) also outline Life's Simple Seven (LS7) criteria, which include core health behaviours (such as smoking, physical activity [PA], diet, and weight) and health factors (such as cholesterol, blood pressure [BP], and glucose control), as main contributors to cardiovascular health (CVH). Over time, Life's Simple Seven (LS7) has evolved into a potent instrument for comprehending strategies to promote healthy ageing within a population. It also may serve as a means to enhance cardiovascular health (CVH) while mitigating the likelihood of developing conditions such as cardiac or cerebrovascular disease, cancer, dementia, and other chronic disorders at the population level. This multifaceted approach contributes significantly to overall well-being (Ioachimescu, O. C., 2022). According to (Ford, Greenlund & Hong, 2012), a 78% lower risk for all-cause mortality and 88% lower risk for mortality from diseases concerning the circulatory system was shown for people who met

more than five criteria from the LS7 health metrics as compared to people who completed less than five.

Why was LE8 introduced

Due to the data accumulated over the past years, quantifying and defining each of the original metrics in LS7 needed calibration for inter and intraindividual variances, refinement and further validation against complex outcomes (Ioachimescu, 2022). Furthermore, some metrics also do not lend themselves to fully continuous quantification scales (for example, smoking was defined categorically with three classifications: never smoked, quit smoking for more than a year, currently smoking) and because some associations of metrics with health were nonlinear, the need to use an ordinal point scoring system for each metric rose (ranging from 0 to 100 points) instead of simple classification as poor (0), intermediate (1), or ideal (2) was suggested (Ioachimescu, 2022). In addition, many researchers suggested an expanded measure of (CVH) that includes sleep as an eighth metric in relation to CVD risk. So, a redefining of the Life Simple Seven (LS7) in Life Essential Eight (LE8) was considered a contributor to enhancing (CVD) primordial and primary prevention efforts. (Makarem et al., 2022).

Why is our research scientifically relevant?

Although intuitively, an approach taking into account the contribution of the various measurements in more detail when arriving at a final score for an individual makes sense, it may also be more cumbersome to collect detailed information. Therefore, to show that LE8 is indeed better in reflecting cardiovascular health than LS7 is needed and currently lacking. We embarked upon that endeavour.

OVERALL AIM

The overall aim is to assess to what extent LE8 predicts CVD better than LS7. The research question is: To what extent does ideal and intermediate cardiovascular health based on the Life Essential Eight approach provide a stronger relation with future cardiovascular events compared to ideal and intermediate cardiovascular health based on the Life Simple Seven approach.

The overall objective was further approached by addressing the following sub-research questions**:** 1) To what extent do CVH metrics individually (CVH factors or predictors) based on LS7 relate to the development of clinical events; 2) To what extent do CVH metrics based on LE8 relate to the development of hazards (rate for the development of clinical event); 3) Comparison of, whether the magnitude of the association with hazards differs between CVH groups based on LS7 and LE8 through NRI (net reclassification benefit).

METHODS

Study population

To perform the analyses, we used the USE-IMT, USE Intima-Media Thickness collaboration (2012) centre data set, an individual participant data set of 16 observational prospective cohorts worldwide and readily available at the University Medical Center Utrecht (USE-IMT, (2012) Utrecht University., 2018). USE-IMT is an ongoing individual participant data meta-analysis (den Ruijter et al., 2012). Eligible participants in our study were identified through data cleaning by taking only complete data for each person from eleven (11) clinics out of sixteen (16). This means the person to be considered for the analysis should have data on all six parameters. Individual information on parameters such as dietary quality and sleep was unavailable. Thus, available baseline data for each participant was set on age, centre, sex, history of CVD, systolic blood pressure (sbp), diastolic blood pressure (dbp), blood pressure, cholesterol levels, body mass index (BMI), PA, exposure to cigarette smoking, and follow-up information on the occurrence of CVD, to build our model of analysis. Baseline characteristics and demographics are given in Table eight (8).

For this study, **the number of participants used for both models was 31.549,** and the number of participants in the USE IMT was originally 62.769.

Measurements - Characteristics extracted from each study- for the LS7 and LE8 model

Descriptive statistics were used to analyze baseline characteristics and demographics, as shown in Table eight (8), together with the number of participants taken into consideration for each metric (which is the same for both models). Table eight (8) also gives the measurements for each metric. When appropriate, characteristics are presented in absolute numbers, percentages or means with standard deviations. In the data set, we had data from young adults, middle-aged women and men, and older people (22 to 90 years old) from the 18 population-based cohort studies.

Per Cohort, we averaged all available measurements (age, systolic blood pressure (sbp), diastolic blood pressure (dbp), blood pressure (BP), cholesterol levels, body mass index (BMI), PA, and exposure to cigarette smoking). This choice was based on the observation that the magnitude of the relation between each metric and the cardiovascular events risk does not differ significantly across various measures for each Cohort. The values were used in the analysis. To account for differences in absolute levels across cohorts because of age and sex differences, we calculated HR by the use of the Cox regression model for each clinic separately based on LS7 and LE8 created by subtracting the individual values from the Cohort for each metric (Figure one (1) and Figure two (2)). First-time CVD (myocardial infarction and first-time stroke) were included as a combined endpoint. These included both fatal and nonfatal events (den-Ruijter et al., 2012). So, the model, as presented in Table 1, was followed.

Table 1: The UMC data set has the following Baseline Risk and Follow-up Characteristics of the Cohorts:

	Based on the Ideal versus Poor health group, Score -HR, % (CI)						
Source and Model	Overall	Men	Women	Follow up, y Median,y	MI, No.	Strokes, No.	First-Time MI or Stroke No.
USE-IMT (LS7)	0.251 (0.207-0.306)	0.359 (0.238-0.480)	0.250 (0.153-0.347)	11 (3.7)	2614	2079	4705
USE-IMT (LE8)	0.237 (0.209-0.269)	0.336 (0.255-0.417)	0.240 (0.178-0.302)				

Abbreviations: IMT- Intima-Media Thickness, MI- myocardial infarctions

Input indicators:

Event follow-up – In our case, the follow-up time is 11 years in USE-IMT with a median of 3.7 years, during which 4705 first-time myocardial infarctions or first-time strokes occurred.

Our study's **event indicator** is the development of a clinical cardiovascular event (myocardial infarction and first-time stroke) in a given time interval (11 years, with a median follow-up of 3.7 years). In the data set, we have defined the event as 0 and 1 (not developing the event and development of the event, respectively). In the data set, the follow-up time for each participant is the period at risk (as defined as the time from baseline assessment to event occurrence, death, loss to follow-up or end of study, whatever comes first), and units of measurement are days. All data were collected at baseline (fixed - or non-time-varying or time-invariant covariates), maintaining the same value throughout the observation period for a given case. They are specified as fixed because they are only measured once, even though they could change over time (Austin, Latouche, & Fine, 2020).

Health groups: Poor, Intermediate and Ideal groups defined for both models.

Output indicators: HRr– hazard rate for developing CVD and HR- hazard ratio for ideal versus poor and intermediate versus poor group for LS7 and LE8 separately were calculated. Stratification per age and sex was also evaluated.

Analysis performed in three steps as follows:

In Step 1, We derived cardiovascular health scores and assessed cardiovascular health status per Life's Simple 7 and Life's Essential 8 per participant. The overall CVH was evaluated by the LS7 score (range from 0-12) and LE8 score (range, 0–100) separately, as well as the score for each factor/metric of physical activity, tobacco/nicotine exposure, body mass index, non–high-density lipoprotein cholesterol, blood glucose, and blood pressure were calculated based on both models. Each metric in LS7 and LE8 consists of poor, intermediate, and ideal strata. For LE8, we took the sum across all six metrics and derived CVH status categories based on thresholds as suggested by Lloyd-Jones et al., (2022b) for CVH assessment. We assessed the overall composite CVH score continuously and then categorically it as poor (< 50),

moderate/intermediate (50– < 80), or ideal/high (80– < 100) based on LE8 components (Makarem et al., 2022).
The classification of poor is from 0 to 4.67, intermediate CVH from 4.67 to 9.34, and significant/ideal more than 9.34 points, according to LS7 (Tsao et al., 2022)
We used metrics and their weights as defined in Tables nine (9) and ten (10) for LS7 and LE8, respectively.

Scoring details for LS7

The total score range was 0 to 14 in the original AHA health metrics as they have seven parameters; however, for this study, it is between 0 and 12 as we analyzed six metrics (Table nine (9)). For each metric (from 0-2) in each clinic, we have given points per patient, which were summed up as a total score per patient depending on the Mean for each metric. The scores of 0–4.67, from 4.67 to 9.34, or bigger than 9.34 points, were regarded as having poor, intermediate, or ideal CVH, respectively (summing the scores for each of the 6 metrics together and dividing the total by 3 (0-2)), (Tsao et al., 2022).

Scoring details for LE8

The scores of 0–100 were used in the original AHA for each metric according to LE8 – fully explained in Table ten (10). For example, an ideal score of 100 points was given for BMI less than 25 kg/m2; for SBP and DBP, 100 points were given when the BP is lower than 120/80, and blood glucose when a patient has no history of diabetes and FBG <100 (or HbA1c <5.7) and non–HDL cholesterol (mg/dL) of less than 130. The LE8 score is the mean value of the six components and is calculated for each individual by summing the scores for each of the six (6) metrics and dividing the total by six (6). In the end, the summing score of 100 is split into three (3) categories, i.e., ≥80: ideal, 50–80: intermediate or <50: poor (Makarem et al., (2022)).

Step 2: We applied Cox regression as a method for analyzing survival data (also known as the continuous-time hazards model, Cox proportional hazard, or even Cox semi-parametric model), where the outcome of interest is the development of clinical cardiovascular disease (CVD) events within a specified time interval, conditional on specific health metrics, with a 95% confidence interval (CI) (Miller, 2008; Su et al., 2022) using R version 4.2.3 (R Core Team, 2023). The calculations performed in R are given in the appendixes as a summary of the analysis performed in R.
Cox regression analysis was performed separately for LS7 and LE8, and calculating the coefficients and the hazard ratios of the covariates, together with the measure of concordance, was completed. In addition, the likelihood ratio, AIC was calculated using R 4.2.3 and bootstrapping to test the internal validity of our model by drawing samples with replacements from the original data set and of the same size as the original one (Koletsi, & Pandis, 2017).

Cox proportional hazard

The calculations were performed in R for each model (LS7 and LE8) using Cox proportional hazards as presented in formula one (1) (Nikulin et al., 2016):

$$\text{Model: } \lambda_i(t) = \lambda_0(t)\, e^{\beta_1 X_{1i} + \cdots + \beta_p X_{pi}} \quad (1)$$

In our case, $X_{1i},\ldots,X_{pi}$ represent the explanatory variables (metrics, factors, predictors, and/or covariates) for an individual (i), and $\beta_1,\ldots,\beta_p$ are the coefficients quantifying their effects. Meanwhile, $e\beta j$ represents the hazard ratio, indicating how many times the hazard increases when the value of X_j increases by 1 unit (Su et al., 2022). Our case demonstrates that a patient within one group, with a specific metric value, can experience a certain percentage reduction or increase in hazard for one unit change in the specific matric compared to someone with poor status, while other parameters remain constant.

Based on the calculation of the Hazard rate, the Hazard Ratio was calculated taking into consideration the health scores per group (poor, intermediate and ideal) for both LS7 and LE8 models separately and as presented in formula two (2) (Nikulin et al., 2016):

$$\lambda_0(t)\, e^{\beta_1 X_{11}+\cdots+\beta_p X_{p1}} \Big/ \lambda_0(t)\, e^{\beta_1 X_{12}+\cdots+\beta_p X_{p2}}$$

$$= e^{\beta_1(X_{11}-X_{12})+\cdots+\beta_p(X_{p1}-X_{p2})} = h_{1vs2} \qquad (2)$$

We have also calculated the changes in the hazard ratio based on one point change for the given metric for both models separately in R 4.2.3.

Model Assumptions and Model Fits

Before reporting the results of Cox regression, we checked the state of model fit and assumptions. The main assumptions for the Cox proportional hazards model are proportional hazards, independence of censoring, and linearity of continuous covariates (Harrell, Jr., & Harrell, 2015). However, as we are applying the Cox model for assessing the probability of event occurrence for each health metric within a given time period for two main models, the assumptions mentioned are irrelevant. Furthermore, nonlinearity is not an issue for categorical variables (ElHafeez et al., 2021), and proportional Hazards are irrelevant since time is not a main exploratory variable in our case, testing the two main models. Besides this fact, we wanted to check if the LE8 model fits better than the LS7, meaning that the predictions made using the LE8 metrics are more precise than those based on LS7 and to re-confirm this using statistical tests in R. We performed the Akaike information criterion, calculating *AIC* for model LS7, and model LE8.

Model diagnostics and validation

We performed model diagnostics and validation to evaluate the quality and robustness of the Cox regression. So, we checked for outliers, influential observations, multicollinearity, and misclassification – noise (Miller, 2008). Regarding influential observations - we can use leverage, Cook's distance, variance inflation factor, and martingale residuals to identify and handle potential problems (Miller, 2008). In our case, we used the following function in R: "*ggcoxdiagnostics function from survminer package"* to Test influential observations.

To validate the model's predictive performance and accuracy using external or internal data, we could use cross-validation, bootstrap, or calibration plots to measure and improve our model's validity (Miller, 2008). We used bootstrapping by drawing samples with replacements

from the original data set and the same size as the original one to test the internal validity (please see the summary of statistics performed in R for the calculations made).

Step 3: We evaluated concordance in CVH status based on LE8 versus LS7, comparing the two methods using the NRI- net reclassification improvement method (index).

In general terms, the NRI measures the improvement in risk prediction when a new marker is added to an existing model. However, it can also compare two models and determine which is better at predicting the risk (Kerr, Wang, Janes, McClelland, Psaty, Pepe, 2014)**.** The net reclassification improvement index (NRI) measures how well a new model reclassifies subjects compared to an old model (Laine et al., 2019). The maximum value for NRI is 2 (Kerr et al., 2014). It can range from -2 to 2, where a value of 2 indicates perfect reclassification, a value of 0 indicates no improvement in reclassification, and a value of -2 indicates perfect misclassification (Kerr, 2022). The NRI is calculated separately for cases and controls, where the NRI Case is the proportion of cases that are correctly reclassified into a higher risk category minus the proportion of cases that are incorrectly reclassified into a lower risk category. The NRI Control is the proportion of controls correctly reclassified into a lower-risk category minus the proportion of controls incorrectly reclassified into a higher-risk category (Kerr et al., (2014)). We used 3+ categorical NRI (for the three categories of CVH – Poor, Intermediate and Ideal) to compare both models. The following formulae were used as per the article from Penecina et al. (2008):

$$\hat{P}(\text{up}|D=1) = \hat{p}_{\text{up,events}} = \frac{\#\text{ events moving up}}{\#\text{ events}}$$

$$\hat{P}(\text{down}|D=1) = \hat{p}_{\text{down,events}} = \frac{\#\text{ events moving down}}{\#\text{ events}}$$

$$\hat{P}(\text{up}|D=0) = \hat{p}_{\text{up,nonevents}} = \frac{\#\text{ nonevents moving up}}{\#\text{ nonevents}}$$

$$\hat{P}(\text{down}|D=0) = \hat{p}_{\text{down,nonevents}} = \frac{\#\text{nonevents moving down}}{\#\text{ nonevents}}$$

The NRI is estimated as

$$\widehat{\text{NRI}} = (\hat{p}_{\text{up,events}} - \hat{p}_{\text{down,events}}) - (\hat{p}_{\text{up,nonevents}} - \hat{p}_{\text{down,nonevents}}) \quad (3\text{-}7)$$

In table five (5)-Total NRI, the rows represent risk categories from the old model and the columns represent risk categories from the new model. Both models place those falling in cells on the diagonal in the same category. In contrast, those above the diagonal are classified as higher risk by the new model and those below the diagonal are classified as having decreased risk by the latest model. Tables six (6) and seven (7) are also given, representing the data for reclassification of Cases/events and Non-Cases/non-events.

Regarding model validation, as we are using the same data set for both models and according to Leening, Vedder, Witteman, Pencina, & Steyerberg (2014), the NRI cannot be affected by miscalibration, meaning that the average predicted risk is not close to the event rate.

RESULTS

The baseline characteristics of the cohorts are presented in Table eight (8). Most of the studied population was white; 53% of the cohorts were women, and the median age was 58. The median follow-up was 3.7 years within 11 years, during which 4007 first-time myocardial infarctions or first-time strokes occurred.

Univariable Hazard ratios by cohorts

In our example, the forest plot shows the point estimates and confidence intervals for eleven (11) out of sixteen (16) cohorts. As can be seen, there is a big variation in results/HR between the clinics and no consistency. The HR are from 0.1 to 1.6 in some clinics based on LS7 and from 0.2 to 2.4 in some cohorts based on the LE8 model. The confidence interval (CI) ranges from 0.1 to 0.3 or 1.1 to 1.2 in another clinic based on LS7 and from 0.1 to 0.5 or from 2.4 to 2.5 in another clinic based on LE8. When analyzing the data per clinic, it has been seen that the variation is due to the age difference as a predicted variable among cohorts. For example, in Cohort Eight (8), the mean age is 49.15; in Cohort One (1), the mean age is 57.5; while in Cohort Three (3), it is 72.39, and in Cohort fifteen (15) is 68.8. so we needed to adjust for age and sex.

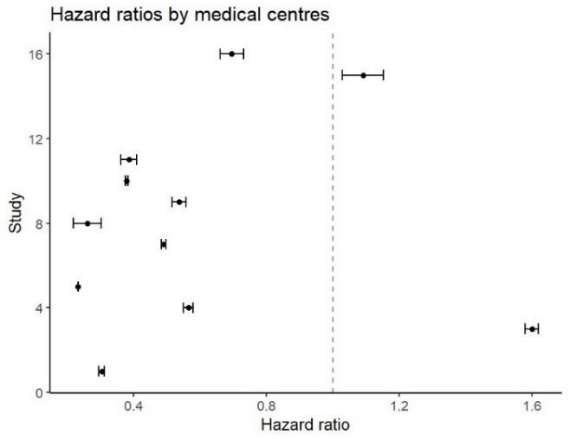

Figure 1: *HR per clinic based on LS7*

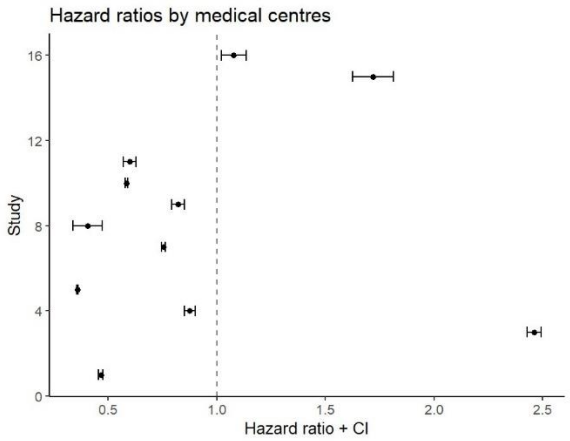

Figure 2: *HR per clinic based on LS8*

Relation between LS7 and LE8 scores to events

The overall estimate of HR for LS7 and LE 8 per health group defined as poor, intermediate and ideal:

- **Table 2:** The HR for both models per health group:

Models	HR		
	0 – poor	1- intermediate (CI)	2 – ideal (CI)
LS7	1 (ref. value for LS7)	0.437 (0.360-0.530)	0.251 (0.207-0.306)
LE8	1 (ref value for LE8)	0.468 (0.416-0.527)	0.237 (0.209-0.269)

Based on LS7, the hazard of the intermediate group is 43.7% with a confidence interval (CI) of (0.360-0.530) relative to the hazard calculated for the poor group (which serves as the reference). This implies that the risk reduction for the intermediate group compared to the poor group is 56.3%. Meanwhile, with LE8, the hazard of the intermediate group is 46.8%, with a CI of (0.416-0.527) relative to that of the poor group. This indicates that the risk reduction for the intermediate group compared to the poor group is 53.2%. Finally, according to LS7, the ideal group has a hazard of 25.1% with a CI of (0.207-0.306) relative to the hazard from the poor group. Similarly, the ideal group in LE8 has a hazard of 23.7% with a CI of (0.209-0.269) relative to the hazard from the poor group. Please refer to Figure 3 for the estimated hazard by model per group:

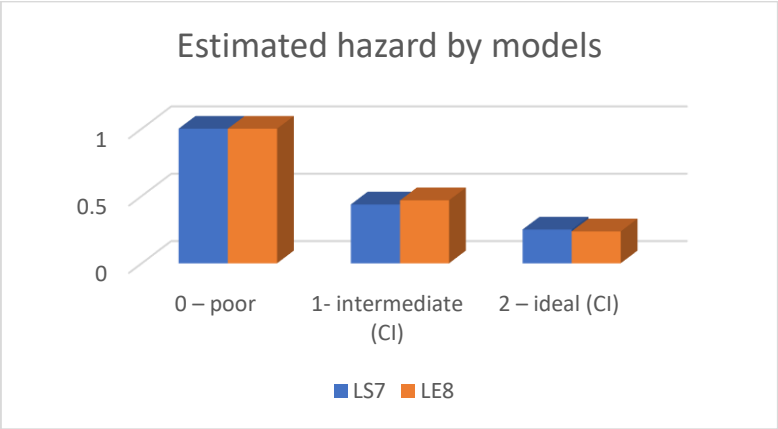

Figure legend: Poor, intermediate or ideal health categories based on the LS7 model (blue bars) Poor, intermediate or ideal health categories based on prediction models' LE8 model (red bars).

Figure 3: *Estimated hazard by models and per group based on the reference value of 0-poor health category*

Model adjusted for age and sex.

So, in order to check for the impact of other factors, such as sex and age, on survival rates and to improve the model accuracy (because of a big variation in results/HR between the clinics and no consistency), we performed stratification analysis in R.

- **Table 3:** The HR for both models for Cohort specific (adjusted models)- sex:

Models	HR=hazard ratios I assume.		
	0 – poor	1- intermediate (CI)	2 – ideal (CI)
LS7			
men	1.438 (1.369-1.507)	0.617 (0.459-0.775)	0.359 (0.238-0.480)
women	1	0.429 (0.302-0.556)	0.250 (0.153-0.347)
LE8			
men	1.397 (1.329-1.466)	0.648 (0.543-0.753)	0.336 (0.255-0.417)
women	1	0.464 (0.383-0.544)	0.240 (0.178-0.302)

As observed, the hazard ratios for men and women within each model exhibit significant differences, while the variations between the models are moderate. The reference parameter is that of women in the poor health group (considered separately for both groups). Specifically, Men in the Poor Group, based on LS7, are 144% more hazardous for cardiovascular disease (CVD) than women. Based on LE8, they are approximately 140% more hazardous for developing CVD. Men in the Ideal Group, compared to women in the poor group and based on LS7, are 35.9% more hazardous. Based on LE8, they are 25% more hazardous. Women in the Ideal Group, compared to women in the poor health group and based on LS7, are 25% more hazardous. Based on LE8, they are 24% more hazardous.

The reference value in our analysis corresponds to the mean covariate within strata. Specifically, all predicted hazard ratios (HRs) are relative to this reference value.

Table 4: The HR for both models for adjusted models)- age-specific:

Models/ age range	HR(mean value)		
	0 – poor	1- intermediate	2 – ideal
LS7			
0- 40 yrs	N/A	0.052	0.029
40-60 yrs	0.654	0.274	0.145
60-80 yrs	2.060	0.981	0.609
>80 yrs	7.412	3.584	2.093
LE8			
0- 40 yrs	0.069	0.057	0.028
40-60 yrs	0.658	0.291	0.138
60-80 yrs	2.172	1.071	0.598
>80 yrs	8.768	3.776	2.058

In our case, the reference individual is approximately 59 years old and belongs to the poor health group (Setting reference levels for Cox regression in R are discussed in [11], [34]). So, the hazard ratio for the age range of 40 - 60 based on LS7 (poor group) is 0.654 relative to the reference value, while for the LE8 (poor group), the HR is 0.658. The hazard ratio for the age range of 40-60 (Ideal Group) based on LS7 is 0.145 relative to the reference value, while for LE8, the HR is 0.138. Notably, between the two models, a significant difference exists in the prediction of HRs for the group over 80 years (7.412 versus 8.768), which belongs to the poor health group.

The results for the coefficients of each metric based on LS7 are as follows:

Similar to logistic regression, the power of beta estimates the hazard ratio (how many times higher or smaller the hazard will be (in one group in comparison to another group in our case) if the predictor increases by one unit (Nikulin et al., 2016). So, the coefficients quantify the size of the effects.

The coefficient for plasma glucose (1) (group intermediate versus poor) is -0.72507, the coefficient for cholesterol one (group intermediate versus poor) is -0.129, the coefficient for physical activity (1) (group intermediate versus poor) is -0.143, the coefficient for smoking (1) (group intermediate versus poor): is -0.076, the coefficient for smoking (2) (group ideal versus poor) is -0.207.

What was expected was that the magnitude of the coefficients of the models would differ as LE8 has more variables and more categories per variable than LS7, whilst the direction would be the same as the coefficients for the variable increase with the risk for both the models. Exceptions arose within the models. For example, one unit change in physical activity showed that the ideal status had a higher hazard rate than the poor group in the LS7. A similar situation arose within the LE8 for the smoking variable, where the above 30 pack /year status has a higher hazard rate than the above 50 pack/year.

The coefficients for LS7 are presented in a forest plot and are given in Figure 4:

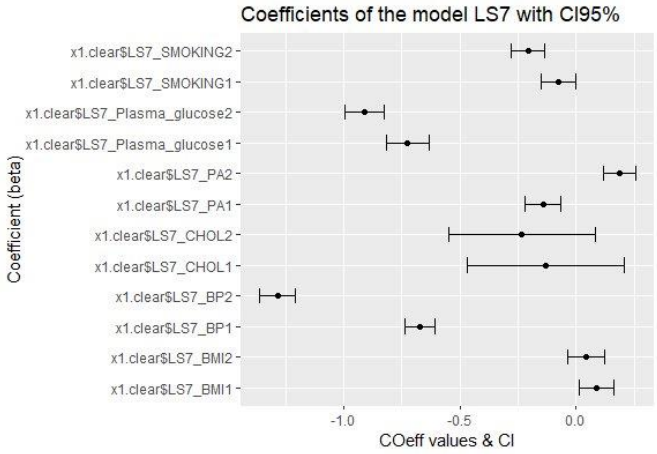

Figure 4*: Coefficients of the model LS7 for six metrics with CI*

The coefficients for each metric based on LE8:

- There are differences for other parameters such as plasma glucose, blood pressure, body mass index, cholesterol and physical activity, as can be seen from figure 5 per health group:

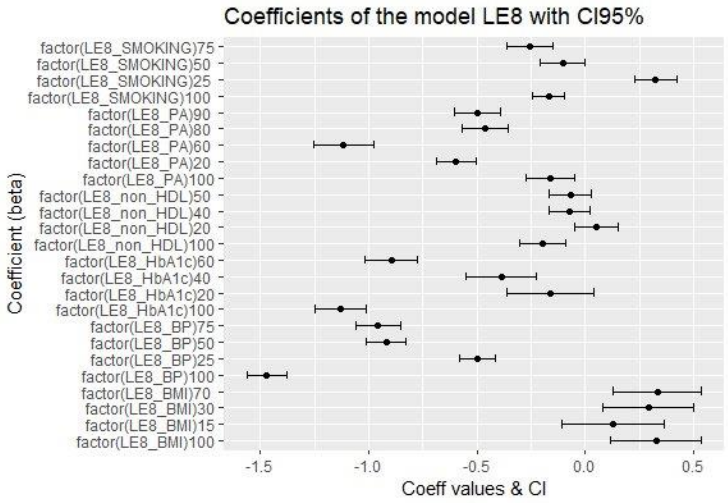

Figure 5*: Coefficients of the model LE8 for six metrics with CI*

The details for the parameters are available in the summary of R calculations for the LE8 model in the appendix.

Regarding the model fit and validation, the model diagnostic parameters have no influence; as for the outliers, we have three groups of data, so there is no data left out that can play the role of outliers in our models. Regarding the influential observations, results show a horizontal line, and the deviance of residuals is almost the same above and below the red line, indicating no influential cases. Also, there is no multicollinearity as the exploratory variables are quite different, so, for example, there is no possibility for colinearity among Smoking and

Physical activity in our case. There is a possibility for misclassification based on typing errors. Still, we can not see a big influence on the final results in our case, which is based on comparing two models using the same data set.

Based on the analysis performed in R, coefficients for LS7 and LE8 for the population, sample and bootstrap models were checked. The results show that bootstrap results are almost equal to the previous models, which confirms the stability and internal validity of the models. For the LS7 model, the coefficients were -0.8277, -1.5502, and -1.5446 for the population, sample, and bootstrap models, respectively. With bootstrapping, the coefficients for LS7 are -1.3795, -2.1464, and -2.1417 for the population, sampling, and bootstrap models, respectively. For the LE8 model, the coefficients were -0.7584, -0.3153, and -0.3069 for the population, sample, and bootstrap models, respectively. The coefficients for LE8 with bootstrapping are -1.4398, -1.1795, and -1.1733 for the population, sampling, and bootstrap models, respectively.

Based on the results, the AIC for LS7 is 91885.83, and for model LE8, it is 91448.99.

Based on NRI-both models compared and expressed with the nr-index:

For NRI analysis, we first used 7900 days as follow-up time. However, according to McKearnan et al., (2018), when we have longer follow-up time, several implications might occur, such as more complexity in the analysis (the risk of the event changes over time so that it can affect the classification of individuals into risk categories, or it can also lead to more missing data due to loss to follow-up). This can result in many censored observations.

So, we faced the following result:

- Total: The total number of observations is 31,549 (out of 62.769 in the USE-IMT data set).
- Cases: The number of cases/events is 4,705.
- Controls/non-events: The number of controls is 4.
- Censored (excluded): The number of censored observations is 26,840.

As can be seen, we have many censored data in our dataset. However, in this situation, events are considered in all cases. If the follow-up time is short, we can have less censored data, but many cases are shifted to a lower-risk group. So, we are dealing with an underestimation of the event results.

According to Kleinbaum and Klein (2012), there are several ways of dealing with censored data:

- We can leave censored data out of the analysis altogether.

- We could assume that the subject “had the event” (= died of CVD event) at the time of censoring.

- We use methods that use the available data, i.e., the information that the subject was alive until the moment of censoring.

Next step in NRI (1) – Shorter follow-up time used (5260 instead of 7900 days)

To have less censored data, we shortened the time of follow-up and performed the following analysis:

From the data, we calculated the so-called " real follow-up time=in which almost all events happened" (for Cases) or had not happened (for non-cases and the censored data). The following graphs have been obtained:

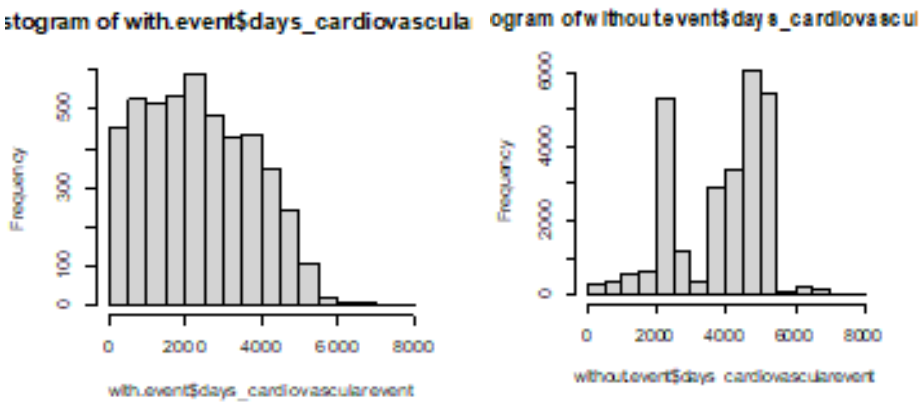

Figure 6: *Graph on the left: Histogram of events for CASES and follow-up time. On the right is the histogram of events for CONTROLS for individuals without events and censored data with follow-up time.*

As can be seen from the right histogram, up to 5260 days of follow-up time (expressed in days), mainly all events happened for cases. Within the same timeframe, we lost most observed data on the controlled and censored side (left histogram).

So we performed the NRI -method on the follow-up time for all observations for up to 5260 instead of 7900 days, and the following results were obtained:

- **Total**: The total number of observations is **31,549** (out of 62.769 in the USE-IMT data set).
- **Cases**: The number of cases/events is **4,638**.
- **Controls/non-events**: The number of controls is **2948.**
- **Censored (excluded)**: The number of censored observations is **23,963**.

As can be seen, the amount of censored data dropped slightly. However, the number of controls improved significantly.

Table 5: Reclassification of data based on both models per group with NRI

	cat_new (LE8)		
cat_orig (LS7)	[0,0.4) =ideal	[0.4,0.7) =int	[0.7,1]=poor
[0,0.4) ideal	*27924*	*1824*	*33*
[0.4,0.7) int	*554*	944	223
[0.7,1] poor	0	33	14

The new model LE8 reclassified 1824 patients from the ideal LS7 health group to the intermediate health group, 33 from the ideal LS7 group to the LE8 poor health group and 223

from the intermediate to the poor group. The model also reclassified 33 patients from the poor LS7 group to the intermediate group in LE8 and 554 from the intermediate LS7 group to the ideal LE8 health group. More participants were re-classified upward to a higher risk CVH group (1824+223+33) than downward to a lower risk group (554 +30).

In addition, the following results are obtained by performing NRI analysis for Cases/events and Non-cases/non/events separately and calculating the NRI for each of these components and total NRI:

- The probability of an upward reclassification in cases is 0.158. (CI 0.129-0.165)
- The probability of a downward reclassification in cases is 0.032. (CI 0.025-0.068)
- The probability of an upward reclassification in controls is 0.042 (CI 0.037-0.041)
- The probability of a downward reclassification in controls is 0.015 (CI 0.0135-0.028)
- NRI Case: The net reclassification improvement (NRI) for cases is 0.125. (CI -0.067-0.134)
- NRI Control: The NRI for controls is -0.027. (CI -0.027-0.012)
- NRI: The overall NRI is 0.098 (CI -0.053-0.107)

The NRI for cases is 0.125, meaning the new model correctly reclassified 12.5% of cases. The NRI for controls is 0.017, meaning the new model correctly reclassified 1.7% of controls. The overall NRI is 0.098, meaning the new model LE8 correctly reclassified 9,8% or approximately 10% of all patients (Kerr, (2022)).

Table 6: Reclassification of data based on both models per group with NRI for cases only

	cat_new (LE8)		
cat_orig (LS7)	[0,0.4)ideal	[0.4,0.7)interim	[0.7,1] poor
[0,0.4) ideal	3465	557	18
[0.4,0.7 interim	133	331	113
[0.7,1] poor	0	16	5

Table 7: Control-Reclassification of data based on both models per group with NRI for controls only

	cat_new (LE8)		
cat_orig (LS7)	[0,0.4) ideal	[0.4,0.7) interim	[0.7,1] poor
[0,0.4) ideal	2747	116	1
[0.4,0.7) interim	47	29	8
[0.7,1] poor	0	0	0

As can be seen from Table six (6), the number of downward (lower risk group) reclassified events/cases is (133+16), while the number of upward (in higher risk group) reclassifications in cases/events is (557+113+18).

Next step in NRI (2)- We can leave censored data out of the analysis altogether.

With the NRI using this method of leaving the censored data out of the analysis, we got the following results: Downward reclassified (lower risk group) events/cases is (64+249), while the number of upward (in higher risk group) reclassifications in cases/events is (54+317) with 1.2% correct reclassification. For controls, the downward reclassification is (160+101), and the upward is (92+143) with 0.9% correct reclassification.

According to the NRI, overall, 2% are correctly reclassified with LE8 compared to LS7, confirming that LE8, although slightly better than LS7.

Next step in NRI (3) - We could assume that the subject "had the event" (=) at the time of censoring.

Suppose the patients had the event downward reclassified (lower risk group). In that case, events/cases are only 65 and 28519 in the high-risk groups according to both scales with negative signs for the reclassification, indicating an incorrect reclassification. For controls, the downward regression is 8, with 2,940 patients in the high-risk group with 0.3% correct reclassification.

According to the overall NRI, 0.04% are correctly reclassified with LE8 compared to LS7.

According to Pepe, Janes, & Li (2014), using NRI p-values in scientific reporting should be halted.

DISCUSSION

Both models show graded relations with risk, and LE8 does better than LS7. Based on the analysis performed with NRI, the net reclassification improvement index (NRI) is 0.0982 (with 5260 days as a real follow-up time), meaning that LE8 classified subjects 10% more accurately than the old model – LS7. The NRI model shows that more cases are reclassified upward (higher risk group), confirming that the new model is better for predicting CVD than the old one. Additionally, the total Net Reclassification Improvement is greater than 0, indicating again that the new model better predicts overall risk than the previous model. According to the Cox analysis, the health behaviour indicators and health factors based on LE8 show slightly different magnitudes of the hazard of CVD development compared to the indicators based on LS7, while their direction is the same. The metrics from LE8 for the intermediate group have a more significant estimate of the probability of the development of CVD (hazards) than the metrics of LS7, which is not the case for the ideal group. Hazard ratios for men and women within each model differ significantly, while they differ moderately between the models. Individually, per predictor, the effect size slightly differs between the two models for the health groups. Based on LS7, the coefficient for plasma glucose (1), for example (group intermediate versus poor), implies that the intermediate health group has a 52% smaller hazard rate compared to the poor health group, even with a one-unit increase in plasma

glucose, while keeping all other parameters constant. The coefficient for plasma glucose (2) (group ideal versus poor) indicates that even with a one-unit increase in plasma glucose, the intermediate group will have a 59.8% smaller hazard rate than someone with poor health status, while holding other parameters constant. While based on LE8, these parameters are 85%, 68%, 41%, 33% for groups scored 20,40,60 and100, respectively, referring to the poor health group (100%) .

Based on the results obtained, the AIC shows that model LE8 fits better than LS7, meaning that the predictions made using the LE8 matrix are more precise than those based on LS7.

Comparison with other studies

Several other studies have also explored risk prediction models, and their results can provide valuable insights such as the study of Dong et al., (2012), the study of Bambs et al., (2011), the study of Folsom et al., (2011) and many others. Almost all of them align with our main findings, confirming that CVH metrics can be used to predict CVD. The study of Dong et al., (2012) confirms that health behaviour indicators and factors severely impact the CVH, emphasizing that a strong relationship was observed between the adjusted hazard ratios for cardiovascular disease and the number of ideal CVH metrics.

The existing risk prediction tools, such as the Framingham Risk Score, lack accuracy in applicability among certain specific populations, in comparison with the new risk prediction models forthcoming to incorporate additional risk factors such as LS7 and LE8 (Hemann et al., (2007).

Practical implications: The LE8 model is better suited to develop more accurate predictions of the development of clinical events such as atherosclerosis or any other CVD, so this can have significant implications for medicine, leading to more effective prevention. The LE8 model can give accurate information to general practitioners, allowing them to guide their patients in making appropriate lifestyle decisions. For example, low-risk or ideal health can indicate no intervention needed, medium-risk or intermediate health suggests lifestyle changes and high-risk or poor health might indicate both lifestyle changes and pharmaceutical intervention.

According to Folsom et al., (2011), timely prevention in individuals with optimal cardiovascular disease (CVD) risk factors and health behaviours can significantly reduce CVD events. This approach serves as a basis for implementing population-wide strategies to promote cardiovascular health (CVH) and discourage the development of CVD risk factors. Furthermore, the LE8 model can be leveraged to estimate population health and recommend public health interventions.

Theoretical implications: It is just now clearly communicated by the NRI index how much better the LE8 model is in predicting CVD than LS7. This is additional justification to the scientific research that the LE8 should do better than LS7, where some metrics do not lend themselves to fully continuous quantification or are nonlinear and should use an ordinal point scoring (ranging from 0 to 100 points) instead of simple classification as poor (0), intermediate (1), or ideal (2) as was also suggested by Ioachimescu, (2022).

Possible limitations

According to Kerr, et al., (2014), " the net reclassification indices do not discriminate/weigh between different types of reclassification — all upward movements in risk categories count the same, as do all downward movements…". It means that the reclassification from poor to intermediate counts the same as that from poor to ideal. This can potentially lead to underestimation of some scientific results and underperformance of the final output. However, this is not the case in our study, as no items are reclassified from 0 to 3 or vice-versa, and the final aim is a comparison between the two models based on the same baseline.

Although we used only six (6) metrics out of seven (7) for the LS7 model and again six (6) out of eight (8) metrics for the LE8 model as defined by AHA, this does not change our findings. Their generalizibility is not in question as, according to Yi Zheng et al., (2023), only a few CVH metrics and related factors can accurately estimate individuals' overall CVH. In addition, when large prospective cohorts are used, the results observed and implications drawn from this study are generalizable to other populations and study settings. So, it is secured that the analysis is applicable for 6 (six) health metrics and results comparable between LS7 and LE8 models. There was no possibility of questioning the external validity.

Directions for future research: Future analysis can also build on exploring the potential applications of the LE8 model in clinical settings and investigating the factors contributing to its superior predictive accuracy.

In conclusion, we used six (6) components of the CVH approach using data from 31,549 participants from 16 cohorts (mean age 57.80 yrs, 53% women). The results obtained show that LS7's hazard ratio of the ideal health group in relation to the poor health group for developing a cardiovascular event was 25.1% with CI (0.207-0.306), while for the intermediate health category versus poor is 43.7% with CI (0.360-0.530). The corresponding hazard ratio based on LE8 was 23.7% with a CI of (0.209-0.269) and 46.8% (0.416-0.527). So, there is strong evidence for the differences in the hazard ratio on the total level for all six (6) components together calculated according to LS7 and LE8 approaches. The health behaviour indicators and health factors on the total level based on LE8 metrics show the different magnitude of development of CVD compared to the indicators based on LS7, while their direction is the same. Equally, there is clear evidence for different magnitudes and even directions of the effects of the hazard for the individual metrics based on different models used. LE8 through NRI (net reclassification benefit) shows that the magnitude of HR differs for 10% of participants and is a preferred model.

TABLES:

Table 8: Baseline characteristics of cohorts

Institution (CENTER)	Men [no]	Women [no]	Age, mean (Range), y	systolic blood pressure (sbp) mean (mmHg) (SD)	diastolic blood pressure (dbp) mean (mmHg) (SD)	Total cholesterol mean (mmol/L) (SD)	HDL Mean (SD) (mmol/L) (SD)	BMI Mean (kg/m^2) (SD)	Plasma glucose Glycemia levels mean (mmol/L) (SD)	HbA1c Mean (%) (SD)	PA (min/per week) (SD)	Smoking (packs per year (one pack -12 boxes) (SD)
Center 1	462	630	56.57 (46-68)	139.17 (19.24)	85.90 (9.73)	6.18 (1.08)	1.34 (0.37)	24.64 (3.74)	5.10 (1.20)	/	2.79. (1.47)	/
Center 3	1583	0	51.17 (65-98)	136.09 (21.42)	/	5.52 (1.01)	1.45 (0.41)	26.74 (4.66)	5.99 (1.88)	/	3.09 (1.39)	/
Center 4	876	0	51.17 (42-61)	132.67 (15.80)	87.98 (10.04)	5.79 (1.04)	1.31 (0.30)	26.51 (3.44)	4.65 (0.82)	/	3 (1.39)	/
Center 5	6055	7910	53.98 (45-64)	120.95 (18.60)	73.64 (11.15)	5.54 (1.08)	1.35 (0.44)	27.63 (5.34)	5.99 (2.16)	/	2.45 (0.79)	/
Centres 7	2045	1597	59.04 (25-82)	142.92 (21.64)	82.95 (12.84)	6.66 (1.26)	1.54 (0.43)	25.60 (3.71)	4.83 (1.10)	/	2.12 (1.24)	/
Center 8	33	0	44.62 (27-58)	129.45 (17.18)	84.73 (10.85)	5.15 (0.94)	1.15 (0.25)	28.77 (3.85)	5.37 (0.70)	/	3.30 (1.59)	/

Center 9	224	227	65.45 (32-86)	138.29 (19.17)	80.65 (11.42)	5.53 (0.89)	1.49 (0.42)	23.14 (2.84)	5.74 (1.30)	/	1.34 (0.41)	26.76 (55.49)
Center 10	3151	3533	62.15 (44-84)	126.62 (21.52)	71.92 (10.27)	5.05 (0.93)	1.32 (0.38)	28.32 (5.45)	5.35 (1.67)	/	3.00 (1.41)	/
Center 11	118	125	68.24 (60-85)	140.05 (20.44)	82.74 (11.19)	5.83 (1.05)	1.46 (0.38)	27.04 (3.55)	6.06 (1.16)	/	3.36 (1.30)	15.69 (19.92)
Center 12	/	/	/	/	/	/	/	/	/	/	/	/
Center 13	/	/	/	/	/	/	/	/	/	/	/	/
Center 14	/	/	/	/	/	/	/	/	/	/	/	/
Center 15	115	118	67.31 (50-94)	140.07 (22.24)	83.07 (12.74)	5.10 (1.00)	1.14 (0.34)	27.33 (4.85)	5.64 (2.74)	/	1.60 (0.75)	/
Centre 16	87	111	58.99 (50-70)	125.07 (15.94)	75.93 (10.24)	5.92 (0.96)	1.41 (0.39)	26.25 (94.14)	5.12 (0.72)	/	0.69 (0.87)	/
Total – .Absolute number or Mean and (SD)	14 749 (no)	16 800 (no)	59.11 (9.88)	128.20 (21.55)	75.11 (12.09)	5.595 (1.15)	1.378 (0.42)	27.21 (5.08)	5.642 (1.90)	5.342 (0.79)	2.606 (1.21)	16.39 (22.43)
Number of participants for each matric for both models (LS7) and (LE8)	14 749	16 800	31549									

Abbreviations: BMI, indicates body mass index; BP, blood pressure; CVH, cardiovascular health; FBG, fasting blood glucose; HbA1c, haemoglobin A1c; HDL, high-density lipoprotein; PA, physical activity; (sdp), systolic blood pressure, (dbp), diastolic blood pressure, (SD), Standard deviation.

Table 9: Definition of metrics and scoring based on LS7

Metric	Quantification of CVH metric: adults (≥20 y of age)
Fasting plasma glucose	
	Ideal
	<100 md/dL
	Intermediate
	100-125 mg/dL or treated to <100 md/dL
	Poor
	≥126 md/dL
Blood pressure	
	Ideal
	SBP <120 / DBP <80 mmHg
	Intermediate
	SBP 120-139 or DBP 80-89 or treated to <120/<80 mmHg
	Poor
	SBP ≥140 or DBP ≥90 mmHg
Total cholesterol	
	Ideal
	<200 mg/dL
	Intermediate
	200-239 mg/dL or treated to <200 mg/dL
	Poor
	≥240 mg/dL
BMI	
	Ideal
	<25 kg/m3
	Intermediate
	25-29.9 kg/m3
	Poor
	≥30 kg/m3
PA	
	‘Ideal
	≥150 min/wk moderate intensity or ≥75min/wk vigorous intensity or ≥150 min/wk intensity + vigorous
	Intermediate
	1-149 min/wk moderate intensity or 1-74 min/wk vigorous intensity or 1-149 min/wk intensity + vigorous
	Poor
	none
Smoking	

	‘Ideal
	Never or quit >1 year
	Intermediate
	Quit <1 year.
	Poor
	Current smoker

Abbreviations: BMI, indicates body mass index; BP, blood pressure; CVH, cardiovascular health; FBG, fasting blood glucose; HbA1c, haemoglobin A1c; HDL, high-density lipoprotein; PA, physical activity; (sdp), systolic blood pressure, (dbp), diastolic blood pressure, (SD), Standard deviation.

Table 10: Definition of matrics and scoring based on LE8

Domain	**CVH metric**	**Method of measurement**	**Quantification of CVH metric: adults (≥20 y of age)**
Health factors & behaviours	BMI	Measurement: Body weight (kilograms) divided by height squared (meters squared)	Metric: BMI (kg/m2)
		Example tools for measurement: Objective measurement of height and weight	Scoring:
			Points Level
			100 <25
			70 25.0–29.9
			30 30.0–34.9
			15 35.0–39.9
			0 ≥40.0
	Blood lipids	Measurement: Plasma total and HDL cholesterol with calculation of non–HDL cholesterol	Metric: Non–HDL cholesterol (mg/dL)
		Example tools for measurement: Fasting or nonfasting blood sample	Scoring:
			Points Level
			100 <130
			60 130–159
			40 160–189
			20 190–219
			0 ≥220
			If the drug-treated level, subtract 20 points.

	Blood glucose	Measurement: FBG or casual HbA1c	Metric: FBG (mg/dL) or HbA1c (%)
		Example tools for measurement: Fasting (FBG, HbA1c) or nonfasting (HbA1c) blood sample	Scoring:
			Points Level
			100 No history of diabetes and FBG <100 (or HbA1c <5.7)
			60 No diabetes and FBG 100–125 (or HbA1c 5.7–6.4) (prediabetes)
			40 Diabetes with HbA1c <7.0
			30 Diabetes with HbA1c 7.0–7.9
			20 Diabetes with HbA1c 8.0–8.9
			10 Diabetes with Hb A1c 9.0–9.9
			0 Diabetes with HbA1c ≥10.0
	BP	Measurement: Appropriately measured systolic and diastolic BPs	Metric: Systolic and diastolic BPs (mm Hg)
		Example tools for measurement: Appropriately sized BP cuff	Scoring:
			Points Level
			100 <120/<80 (optimal)
			75 120–129/<80 (elevated)
			50 130–139 or 80–89 (stage 1 hypertension)
			25 140–159 or 90–99
			0 ≥160 or ≥100
			Subtract 20 points if treated level.
	PA	Measurement: Self-reported minutes of moderate or vigorous PA per week. Example: Tool for measurement-Questionnaire.	Metric: Minutes of moderate or greater intensity activity per week

			Scoring: Points Minutes
			100 ≥150
			90 120–149
			80 90–119
			60 60–89
			40 30-59
			20 1-29
			0 0
		Measurement: Self-reported use or inhaled Example: Tool for measurement – Self-reported questionnaire	Metric: Combustible tobacco use or second-hand smoke exposure Scoring:
			Points Status
			100 Never smoker
			75 Former smoker
			quite ≥5y
			50 Former smoker
			Quite 1-<5y
			25 Former smoker
			Quite 1
			Inhaled NDS
			0 Current smoker Subtract 20 for leaving with an active indoor smoker in the home.

Abbreviations: BMI, indicates body mass index; BP, blood pressure; CVH, cardiovascular health; FBG, fasting blood glucose; HbA1c, haemoglobin A1c; HDL, high-density lipoprotein; PA, physical activity; (sdp), systolic blood pressure, (dbp), diastolic blood pressure, (SD), Standard deviation.

APENDEXES

1. **Summary R statistics**
2. **Summary R statistics for the NRI model**

REFERENCES

1. Abd ElHafeez, S., D'Arrigo, G., Leonardis, D., Fusaro, M., Tripepi, G., & Roumeliotis, S. (2021). Methods to analyze time-to-event data: the Cox regression analysis. *Oxidative Medicine and Cellular Longevity*, *2021*, 1-6.
2. Amgen (Europe) GmbH (2021). Amgen (Europe) GmbH report 2021 on Secondary prevention of heart attack and stroke, Country profile for the Netherlands. Retrieved from https://www.healthpolicypartnership.com/app/uploads/Secondary-prevention-of-heart-attack-and-stroke-in-Europe-Netherlands.pdf.
3. Austin, P. C., Latouche, A., & Fine, J. P. (2020). A review of the use of time-varying covariates in the Fine-Gray subdistribution hazard competing risk regression model. *Statistics in medicine*, *39*(2), 103-113.
4. Bambs, C., Kip, K. E., Dinga, A., Mulukutla, S. R., Aiyer, A. N., & Reis, S. E. (2011). Low prevalence of "ideal cardiovascular health" in a community-based population: the heart strategies concentrating on risk evaluation (Heart SCORE) study. Circulation, 123(8), 850-857.
5. Coronado, F., Melvin, S. C., Bell, R. A., & Zhao, G. (2022). Peer Reviewed: Global Responses to Prevent, Manage, and Control Cardiovascular Diseases. Preventing Chronic Disease, 19.
6. David, G. K., & Mitchel, K. (2012). Survival analysis: a Self-Learning text. Spinger.
7. Den Ruijter, H. M., Peters, S. A., Anderson, T. J., Britton, A. R., Dekker, J. M., Eijkemans, M. J., ... & Bots, M. L. (2012). Common carotid intima-media thickness measurements in cardiovascular risk prediction: a meta-analysis. Jama, 308(8), 796-803.
8. Dong, C., Rundek, T., Wright, C. B., Anwar, Z., Elkind, M. S., & Sacco, R. L. (2012). Ideal cardiovascular health predicts lower risks of myocardial infarction, stroke, and vascular death across whites, blacks, and hispanics: the northern Manhattan study. Circulation, 125(24), 2975-2984.
9. Folsom, A. R., Yatsuya, H., Nettleton, J. A., Lutsey, P. L., Cushman, M., Rosamond, W. D., & ARIC Study Investigators. (2011). Community prevalence of ideal cardiovascular health, by the American Heart Association definition, and relationship with cardiovascular disease incidence. Journal of the American college of cardiology, 57(16), 1690-1696.
10. Ford, E. S., Greenlund, K. J., & Hong, Y. (2012). Ideal cardiovascular health and mortality from all causes and diseases of the circulatory system among adults in the United States. circulation, 125(8), 987-995.
11. H. J. Motulsky, " Setting reference levels for Cox regression" , GraphPad StatisticsGuide. Accessed 10 January 2024. https://www.graphpad.com/guides/prism/latest/statistics/stat_cox_regression_reference_level_tab.htm
12. Harrell, Jr, F. E., & Harrell, F. E. (2015). Ordinal logistic regression. *Regression modelling strategies: with applications to linear models, logistic and ordinal regression, and survival analysis*, 311-325.

13. Hemann, B. A., Bimson, W. F., & Taylor, A. J. (2007). The Framingham Risk Score: an appraisal of its benefits and limitations. American Heart Hospital Journal, 5(2), 91-96. H. J. Motulsky, " Setting reference levels for Cox regression" , GraphPad StatisticsGuide. Accessed 10 January 2024.
14. Herrington, W., Lacey, B., Sherliker, P., Armitage, J., & Lewington, S. (2016). Epidemiology of atherosclerosis and the potential to reduce the global burden of atherothrombotic disease. Circulation research, 118(4), 535-546. https://www.graphpad.com/guides/prism/latest/statistics/stat_cox_regression_reference_level_tab.htm
15. Ioachimescu, O. C. (2022). From Seven Sweethearts to Life Begins at Eight Thirty: A Journey From Life's Simple 7 to Life's Essential 8 and Beyond—Journal of the American Heart Association, 11(21), e027658.
16. Kerr, K. F., Wang, Z., Janes, H., McClelland, R. L., Psaty, B. M., & Pepe, M. S. (2014). Net reclassification indices for evaluating risk-prediction instruments: a critical review. *Epidemiology (Cambridge, Mass.)*, *25*(1), 114.
17. Kerr, K. F. (2022). Net Reclassification Index Statistics Do Not Help Assess New Risk Models. *Radiology*, *306*(3), e222343.
18. Koletsi, D., & Pandis, N. (2017). Survival analysis, part 3: Cox regression. *American journal of orthodontics and dentofacial orthopedics*, *152*(5), 722-723.
19. Koopman, C., Vaartjes, I., Blokstra, A., Verschuren, W. M. M., Visser, M., Deeg, D. J. H., ... & Van Dis, I. (2016). Trends in risk factors for coronary heart disease in the Netherlands. BMC Public Health, 16, 1-8.
20. Leening, M. J., Vedder, M. M., Witteman, J. C., Pencina, M. J., & Steyerberg, E. W. (2014). Net reclassification improvement: computation, interpretation, and controversies: a literature review and clinician's guide. Annals of internal medicine, 160(2), 122-131
21. Lopez, E. O., Ballard, B. D., & Jan, A. (2023). Cardiovascular disease. In *StatPearls [Internet]*. StatPearls Publishing.
22. Lloyd-Jones, D. M., Allen, N. B., Anderson, C. A., Black, T., Brewer, L. C., Foraker, R. E., ... & American Heart Association. (2022). Life's essential 8: updating and enhancing the American Heart Association's construct of cardiovascular health: a presidential advisory from the American Heart Association. *Circulation*, *146*(5), e18-e43.
23. Lloyd-Jones, D. M., Ning, H., Labarthe, D., Brewer, L., Sharma, G., Rosamond, W., ... & Perak, A. M. (2022a). Status of cardiovascular health in US adults and children using the American Heart Association's new "Life's Essential 8" metrics: prevalence estimates from the National Health and Nutrition Examination Survey (NHANES), 2013 through 2018. Circulation, 146(11), 822-835.
24. Lloyd-Jones, D. M., Ning, H., Labarthe, D., Brewer, L., Sharma, G., Rosamond, W., ... & Perak, A. M. (2022b). Status of cardiovascular health in US adults and children using the American Heart Association's new "Life's Essential 8" metrics: prevalence estimates from the National Health and Nutrition Examination Survey (NHANES), 2013 through 2018. Circulation, 146(11), 822-835.
25. Makarem, N., Castro-Diehl, C., St-Onge, M. P., Redline, S., Shea, S., Lloyd-Jones, D., ... & Aggarwal, B. (2022). Redefining cardiovascular health to include sleep: prospective

associations with cardiovascular disease in the MESA sleep study. *Journal of the American Heart Association*, *11*(21), e025252.
26. McKearnan, S. B., Wolfson, J., Vock, D. M., Vazquez-Benitez, G., & O'Connor, P. J. (2018). Performance of the net reclassification improvement for nonnested models and a novel percentile-based alternative. American journal of epidemiology, 187(6), 1327-1335.
27. Miller, J. E. (2008). Writing about Hazards Models: Practical Guidelines for Effective Presentation Jane E. Miller, Ph. D January 2008.
28. Nikulin, M., & Wu, H. D. I. (2016). The Cox model and its applications. Berlin: Springer.
29. Pahwa R, Jialal I. Atherosclerosis. 2023 Aug 8. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2023 Jan–.PMID: 29939576.
30. Pencina, M. J., D'Agostino Sr, R. B., D'Agostino Jr, R. B., & Vasan, R. S. (2008). Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Statistics in medicine, 27(2), 157-172.
31. Pepe, M. S., Janes, H., & Li, C. I. (2014). Net risk reclassification p values: valid or misleading?. Journal of the National Cancer Institute, 106(4), dju041.
32. R Core Team (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
33. Su, P. F., Lin, C. C. K., Hung, J. Y., & Lee, J. S. (2022). The Proper Use and Reporting of Survival Analysis and Cox Regression. *World Neurosurgery*, *161*, 303-309.
34. "Survival analysis: Predictions of a Cox Model", R-manual, Documentation for package 'survival' version 3.5-7. Accessed 10 January 2024. https://stat.ethz.ch/R-manual/R-devel/library/survival/html/predict.coxph.html
35. Thomas LE, O'Brien EC, Piccini JP, D'Agostino RB, Pencina MJ. Application of net reclassification index to non-nested and point-based risk prediction models: a review. European Heart Journal. 2019 Jun;40(23):1880-1887. DOI: 10.1093/eurheartj/ehy345. PMID: 29955849; PMCID: PMC6568208.
36. Tsao, C. W., Aday, A. W., Almarzooq, Z. I., Alonso, A., Beaton, A. Z., Bittencourt, M. S., ... & American Heart Association Council on Epidemiology and Prevention Statistics Committee and Stroke Statistics Subcommittee. (2022). Heart disease and stroke statistics.
37. Utrecht University. (2012). USE-IMT [Data set]. Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht. https://juliuscentrum.umcutrecht.nl/en/studies-en-cohorten/cohort-use-imt
38. World Health Organization (2021). The World Health Organization annual report 2021. Retrieved from https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
39. Y. Koop, R.H. Wimmers , M.L. Bots , I. Vaartjes (2023), *Basic figures for ischemic heart disease,* Retrieved (12,2023) from https://www.hartenvaatcijfers.nl/jaarcijfers/basiscijfers-ischemische-hartziekten-1dd64.