Overestimation of Cardiovascular Mortality Risk by Kaplan-Meier in Competing Risks Settings: A Web-Based Calculator and NHANES Analysis

Mohammad Zaino*

Faculty of Nursing and Health Sciences, Department of Physical Therapy, Jazan University, Jazan, KSA, Saudi Arabia

Abstract: *Background*: Traditional Kaplan-Meier (KM) event rates are widely used for cardiovascular risk prediction and tend to overestimate absolute event risk for patients by censoring competing events, such as non-cardiovascular death. Competing risks analysis (CRA), which account for such terminal events, offers more accurate estimates. However, its application in a web-based health analytics remains limited.

Methods: Using a simulated cohort (n = 2,500; 100 repetitions) and the 1999–2000 NHANES cohort (n = 2,480) with 2019 National Death Index mortality linkage, the researcher compared KM estimates to CRA's Cumulative Incidence Function (CIF), implemented via Aalen-Johansen estimators and Fine-Gray subdistribution hazard models. We assessed relative differences (bias) in 5-, 10-, 15-, and 20-year cardiovascular mortality predictions across risk strata. Findings informed a web-based calculator prototype that dynamically estimates age-specific KM and CIF probabilities while highlighting potential misclassification risks.

Results: KM consistently overestimated cardiovascular mortality risk compared to CIF. In the NHANES cohort, KM estimated the 5-year risk to be 5.85% higher than the actual rate (4.37% vs. 4.13%) and 20-year risk by 28.3% (20.02% vs. 15.60%). In the simulated data, KM overestimated the 5-year risk by 7.63% (5.84% vs. 5.42%) and the 20-year risk by 31.17% (21.37% vs. 16.25%). KM-based models tend to misclassify a substantial portion of patients into higher-risk groups compared to CIF-adjusted models.

Conclusion: This study demonstrates that Kaplan-Meier consistently overestimates cardiovascular mortality in comparison to competing risk methods across five time points, through using both simulated and nationally representative data. We quantify this overestimation and provide an online calculator that shows differences by age. Our tool improves the usability and interpretability of competing risks analysis for older adults in digital health settings, in contrast to tools like SCORE2.

Keywords: Competing risks, Cardiovascular mortality, Web-based health analytics, NHANES, Clinical decision support.

INTRODUCTION

Clinical Importance of CVD Risk Prediction

Cardiovascular disease (CVD) remains the leading cause of death worldwide, leading to 17.9 million deaths annually — about 32% of all deaths worldwide. Making well-informed clinical decisions and putting public health initiatives into action to mitigate this problem depend on having reliable risk prediction tools [1], [2].

Limitations of Kaplan-Meier (KM)

A common method in traditional survival analysis, particularly in CVD studies, for evaluating time-to-event outcomes is the Kaplan-Meier (KM) estimator. When it comes to competing risks, such as non-cardiovascular deaths, which can impede the occurrence of the primary event of interest - cardiovascular death - KM techniques are inadequate. Relative difference (bias) risk estimates may result from this flaw, especially in groups where competing events are common [3].

The KM estimator states even censored individuals (including those with competing events) still have a chance of experiencing the main event of the study. This leads to inflated incidence rates (1). A similar issue arises in heart failure studies with cardiovascular mortality where KM-based estimates of death risk can be off by as much as 6-8% making them less accurate than methods that consider competing risks [4[. This misrepresentation is critical when it comes to risk stratification and treatment allocation in the context of prognostic modelling on digital health platforms. Overestimating a "single" event can lead to unnecessary interventions or misallocation of resources [5].

Competing Risks as a Solution (Fine & Gray, 1999)

A stronger framework has been proposed which will incorporate Cumulative Incidence Functions (CIFs) and Fine-Gray Sub distribution Hazard Models. They

*Address correspondence to this author at the Department of Physical Therapy, Faculty of Nursing and Health Sciences, Jazan University, Jazan, Saudi Arabia, Al Maarefah Rd, Jazan, Postal Code: 45142, P.O. Box: 114, Saudi Arabia; Tel: +966595779417; E-mail: mzaino@jazanu.edu.sa

E-ISSN: 1929-6029/25

concentrate on competing events without depending on independence assumptions. The CIF ensures a more accurate estimation of the absolute risk of an event by competing risks and in turn better prediction of prognosis. For example, when comparing KM and CIF estimates among cohorts of cardiovascular disease (CVD), KM was observed to overestimate the 5-year cardiac mortality rate by 6.2% (43.0% compared to 36.8%) [6]. Furthermore, results from the SCORE2 risk tool - identified as a competing risk - indicate that it may be more effective than traditional aging models such as Framingham in adjusting for non-CVD deaths, especially for younger and high-risk subjects [7].

Heart failure and coronary heart disease are just two of the conditions that fall under the umbrella term "CVD." The evaluation of these conditions can be significantly complicated by the existence of competing risks. For example, the risk of dying from non-cardiovascular causes frequently exceeds the risk of dying from heart-related causes in patients with heart failure. This demonstrates how important it is to weigh these conflicting risks in order to make the appropriate diagnosis [8]. Likewise, for stroke survivors, we must take into account the risk of dying from other causes, such as cancer, in addition to the risk of having another stroke [9].

Gap in Web-Based Tools and Study Aim

The interactive tools health analytics in this context are used to support the process of identifying risk by applying data and risk-stratification at an individual level. The majority of these platforms, however, continue to use KM-based estimates, which will be between 14% and 20% of high-risk patients [10]. Misdiagnosis can potentially lead to overtreatment or treatment that affects healthcare resources and expenses for patients. Competing risk approach has to be included in such platforms for continued accuracy and clinical decision support.

The goal of this research, as well as the seemingly evident gaps in the past research, is founded on several core issues, specifically because there is more emphasis on competition risk:

- Test for relative bias: In order to accurately predict risks, we need comprehensive studies comparing Kaplan-Meier (KM) and Competing Risk Analysis (CRA) methods, specifically in modelled cardiovascular cases where relative differences (bias) could be a problem.
- Implications for digital health device recommendations: There is a need to identify and examine the ways in which such biases (differences) affect digital health devices, particularly among older individuals and those with complicated conditions.
- Developmental Recommendations: We must advise how to enable the incorporation of CRA into real-time risk prediction models with a goal of enhancing precision and dependability in any web-based health analysis. By filling these loopholes, this research will enhance the accuracy of digital platforms health tools and enable clinical practice in managing cardiovascular diseases.

Future Directions in the Study of Competing Risks

Rapid advancements in the field of competing hazards are providing new methods for evaluating dynamic risk factors and tailoring treatment for upcoming research. We can improve the predictive accuracy of competing risk models for individuals by combining lifestyle, genetic, and biomarker data. Furthermore, our risk estimates can be further refined by applying machine learning techniques to reveal intricate relationships between risk factors and competing events [11].

Integration with Modern Risk Prediction Tools

Novel opportunities for adapting to competing risks are presented by recent developments in cardiovascular disease (CVD) risk prediction, especially using machine learning models. For example, machine learning-based models that used 75 different inputs showed better discrimination (AUC 0.74) than the WHO risk charts (AUC 0.51) in a study on a Sri Lankan cohort. The capacity of machine learning to take competing risks into account contributes to this improvement [12]. It's crucial to remember that there are population differences; for instance, in rural India, Globe-risk may perform better than WHO and Australian risk scores when the model is based on clinical criteria [13]. These results highlight the necessity of calibrating predictive algorithms in order to better handle competing risks and adjust to particular situations.

METHODS

NHANES Data Extraction and Preparation

The 2019 mortality follow-up dataset from the National Center for Health Statistics (NCHS) was linked

to publicly available data from the 1999–2010 cycles of the National Health and Nutrition Examination Survey (NHANES) [14]. Real-world data has been used to validate the simulation results. Participants were defined as those who were 45 years of age or older at baseline with average age 64.14 years old. Mortality follow-up was available until December 31, 2019, with a follow-up period of 20 years allowed.

The NHANES Demographics Files (DEMO) was retrieved for baseline and demographic data for each of the six survey cycles. The read_xpt() function in R was used to load the NHANES demographic files [15] listed below:

```
r
# Load the real NHANES demographic files
demo_9900 <- read_xpt("DEMO.XPT")     # 1999-2000
demo_0102 <- read_xpt("DEMO_B.XPT")   # 2001-2002
demo_0304 <- read_xpt("DEMO_C.XPT")   # 2003-2004
demo_0506 <- read_xpt("DEMO_D.XPT")   # 2005-2006
demo_0708 <- read_xpt("DEMO_E.XPT")   # 2007-2008
demo_0910 <- read_xpt("DEMO_F.XPT")   # 2009-2010
```

These datasets were combined with the associated mortality file (NHANES_1999_2000_MORT_2019_ PUBLIC. dat) and pertinent examination datasets after being harmonized to guarantee consistent variable coding across cycles. SEQN, a unique respondent identifier, was used to merge the data. The resulting dataset served as the foundation for the survival analysis that contrasted competing risks and Kaplan-Meier methods for estimating cardiovascular mortality [16].

From the date of the baseline interview to the date of death or censorship in 2019, the follow-up period was computed. The ICD-10 underlying cause codes from the mortality file were used to classify the cause of death; ICD-10 codes I00–I99 (diseases of the circulatory system) were used to define cardiac death [17].

After keeping participants who had eligstat = 1 (eligible for mortality follow-up), we calculated the survival time (time) by converting the number of months from the interview to death or censoring to years (permth_exm / 12). Participants who were alive at the end of follow-up or lost to follow-up were considered administratively censored.

By assessing the effect of competing risks on long-term risk estimates, the analyses were intended to assist in the creation of precise risk prediction models for web-based health analytics, such as online cardiovascular risk calculators.

Covariates such as comorbidities were not included in the survival analysis, as the primary aim was to isolate differences between Kaplan-Meier and competing risks methods in a structure consistent with the simulation study.

Participants with incomplete survival time or cause of death information were excluded. To maintain the natural disease mix and support the analysis's estimator-focused goal, eligibility was limited to individuals with eligstat = 1 and no exclusion was based on pre-existing cardiovascular disease.

Simulation Study Design

A simulation of a synthetic cohort of 2500 individuals (repeated 100 times) is conducted with ages sampled from a uniform distribution between 45 and 85 years. Cardiac and non-cardiac death times were simulated using age-dependent exponential hazard functions. The exponential hazard function was chosen because it is straightforward and simple to understand in simulation-based comparisons of estimation techniques, even though Weibull and Gompertz distributions might more accurately represent actual aging-related mortality processes. This presumption prevents over parameterization and enables consistent, linear age effects across simulations. Previous simulation studies that looked at estimator bias have used similar simplifications [18, 19].

The cause-specific hazard rate for non-cardiac death was modelled as:

$$\lambda_{NC}(t|\, age) = \left[0.04 + 0.08\times\left(\frac{age}{85}\right)\right]\times 0.30$$

The term $\lambda_{NC}(t|\, age)$ represents the instantaneous hazard rate at time *t* for non-cardiac death in individuals of a given age. That means the risk of experiencing the event (e.g., non-cardiac death) at time *t*, conditional on survival up to *t* and the individual's age. This linear specification of age: age_effect = 0.08× (age / 85) was chosen based on preliminary analyses of the NHANES 1999–2000 cohort, where the effect of age on cardiac death hazard was evaluated for linearity (see statistical analysis section).

The age-independent component (or "intercept term") of the hazard function (0.04) corresponds to the minimum risk at younger ages (e.g., 45 years).

If we for example consider an individual for a 60-year-old

$$\lambda_{\mathrm{NC}}(t|\,60) = \left[0.04 + 0.08\times\left(\frac{60}{85}\right)\right]\times 0.30 = [0.04 + 0.0565]\times 0.30 = 0.02895$$

The hazard rate is then 2.9% per unit time (e.g., per year). This represents the instantaneous risk of non-cardiac death for a 60-year-old, assuming they've survived up to time t.

This formulation accounts for the increasing risk of non-cardiac mortality with aging, by scaling the base hazard using a linear function of age. The aforementioned term refers to an age-dependent hazard rate, normalized to a maximum age of 85 years, which rises in direct proportion to age. The entire expression is multiplied by a calibration factor of 0.30 to lower the non-cardiac death rate in relation to the total event rate. This scaling accounts for the assumption that non-cardiac death, while influenced by age, plays a smaller role than cardiac death in the overall mortality of this synthetic cohort. This factor was selected heuristically to represent a plausible imbalance between competing risks and to make estimation bias easier to interpret, rather than being empirically derived from external datasets. We recognize that the absolute event rates might not correspond to those seen in real-world cohorts, even though this decision facilitates a controlled comparison between Kaplan-Meier and competing risks approaches. However, this assumption has no effect on the internal validity of estimator comparisons.

By adjusting this multiplier, the model maintains a realistic balance between competing risks, allowing for clearer comparison of estimation bias between Kaplan-Meier and competing risks methods. Without this factor, the initial hazard rates derived from the baseline and age terms might have resulted in an unrealistic number of non-cardiac deaths compared to cardiac deaths in the simulation.

The cardiac death hazard was defined as a proportion of the non-cardiac hazard:

$$\lambda_C(t|\,age) = 0.40\times\lambda_{NC}(t|\,age)$$

This method was applied to the simulated data in order to preserve a consistent and comprehensible structure of competing risks. In order to ensure that cardiac events are less common than non-cardiac events, the simulation introduces a controlled level of competition between causes of death by modelling cardiac mortality as 40% of the non-cardiac hazard across all ages.

The simulation is made simpler by this proportional relationship while maintaining a realistic risk disparity, which is frequently seen in older populations where non-cardiac causes (such as cancer and respiratory diseases) frequently account for the majority of deaths. More importantly, this structure allows for clear evaluation of bias (differences) and misclassification in Kaplan-Meier estimates, which assume independence of competing events, versus methods that properly account for competing risks like the cumulative incidence function (CIF) [20].

The event time for each person was calculated as the lowest of the administrative censoring time, non-cardiac death time, and cardiac death time. A uniform distribution of administrative censoring times was sampled:

$$t_{censor} \sim Uniform(17, 21)$$

This range was chosen to represent the actual follow-up duration seen in the NHANES dataset, where the longest follow-up period was just over 20 years (up to about 20.57 years). The external validity of the simulation results was improved by matching the censoring distribution with the empirical data, which guaranteed that the simulated risk estimates were assessed over a similar time horizon [21].

The event type was assigned based on which time was earliest:

- If $t_{censor} < \min(t_{noncardiac}, t_{cardiac})$: censored (event =0)
- If $t_{cardiac} < t_{noncardiac}$ and $t_{cardiac} < t_{censor}$: cardiac death (event =1)
- If $t_{noncardiac} < t_{cardiac}$ and $t_{noncardiac} < t_{censor}$: non-cardiac death (event =2)

The simulation was designed to generate an outcome distribution reflective of NHANES data, ensuring its suitability for comparing Kaplan-Meier (KM) and cumulative incidence function (CIF) estimates.

Additionally, the number of individuals at risk at 5, 10, 15 and 20 time point was evaluated to assess the precision of long-term estimates.

STATISTICAL ANALYSIS

The statistical methods in this study were applied to evaluate and enhance the accuracy of risk estimates

for potential integration into web-based health analytics platforms, where precise predictions are essential for user-facing tools like digital risk assessment applications (e.g., risk calculators, patient-facing apps).

To evaluate whether the effect of age on the hazard of cardiac death was appropriately modelled as linear, we fitted Cox proportional hazards models including age as a linear term, a quadratic term, and as a natural cubic spline (with 3 degrees of freedom) to the NHANES cohort. Model fit was compared using likelihood ratio tests and Akaike Information Criterion (AIC), with statistical significance assessed at the 0.05 level.

Kaplan-Meier Estimation

For the NHANES dataset, the Kaplan–Meier (KM) estimator is used to calculate the cumulative incidence of cardiac death, treating non-cardiac deaths as censored.

The KM estimator or survival probability is given as

$$\hat{S}(t) = \prod_{t_i<t}(1-\frac{d_i}{n_i}),$$

with warnings when competing risks exist (10).

d_i : Represents subjects experiencing the event of interest (e.g., cardiovascular death) at time t_i . That is if 3 patients die from cardiac arrest at 12 months follow-up then d_i=3. and t_i denotes subjects alive and uncensored just before the time t_i.

The same approach was applied to the simulated dataset to ensure comparability. While standard, this method does not account for the informative nature of competing events, potentially overestimating the risk of cardiac death.

Cumulative Incidence Function (CIF)

To obtain estimation of the cumulative incidence of cardiac death while properly accounting for the competing risk of non-cardiac death, the Aalen-Johansen estimator using the **cuminc**() function from the cmprsk R package was used. The Aalen-Johansen estimator is a non-parametric method specifically used to estimate the Cumulative Incidence Function (CIF) in the presence of competing risks.

The CIF estimates the probability of experiencing a specific event type (e.g., cardiovascular death) by time t, while accounting for competing risks (e.g., non-cardiovascular death). Its formula is:

$$\mathrm{CIF}_k(t) \int_0^t \lambda_k(s).\hat{S}(\bar{s})\, ds\,,$$

where $\lambda_k(s)$ is the instantaneous rate (cause-specific hazard rate) of event k at time s conditional on surviving all events (including competing risks) up to s.

$\hat{S}(\bar{s})$ is the overall survival function at time $\bar{s}$. This represents the probability of surviving up to just before time s without experiencing any of the competing events.

This method was used for both NHANES and simulated datasets, directly estimating the probability of cardiac death in the presence of competing risks.

Fine-gray Regression Model

The Fine-Gray model is employed in its baseline form (without covariates) to estimate the sub-distribution hazard of cardiac death for both NHANES and simulated datasets. This model corresponds directly to the cumulative incidence function (CIF) in the presence of competing risks.

Although the Fine-Gray model is typically used for regression with covariates, in this study it was employed solely for estimating the sub-distribution hazard of cardiac death, allowing corroborating the Aalen-Johansen CIF estimates, ensuring consistency in competing risk assessments for web-based applications.

Difference Validation

The bias between methods was evaluated as the absolute difference between KM and CIF estimates at 5 years and computed relative bias as:

$$\text{Relative Diff. (\%)} = \frac{KM - CIF}{CIF} \times 100$$

For clinical interpretation, the absolute difference (KM – CIF) was also reported, representing the overestimation in percentage points.

The external validation was done using the 1999-2000 cohorts of the National Health and Nutrition Examination Survey (NHANES), linked with the mortality follow-up data via the 2019 National Death Index (CDC, 2022; NCHS, 2021). This dataset furthers its evidence base by comparing real-world data and providing benchmarks against the simulated estimates. NHANES is a nationally representative survey widely used in epidemiological research, including studies of cardiovascular risk and mortality that incorporate

competing risks frameworks (e.g., [21, 22]). NHANES validation provided real-world benchmarks for competing risks approaches, enhancing their applicability to web-based health analytics platforms that require accurate risk predictions.

Visualization

Cumulative incidence curves from both the KM and CIF estimators were plotted to visually demonstrate the degree of overestimation attributable to ignoring competing risks. The vertical axis of the plots was standardized (0 to 0.25) to accommodate both NHANES and simulation KM and CIF estimates, facilitating visual comparisons for web-based risk visualization tools.

Software

All analyses were conducted using R (version 4.3.1). The survival package (version 3.5-5) was used for Kaplan-Meier estimation, and the cmprsk package (version 2.2-11) was used for Cumulative Incidence Function (CIF) estimation.

Prototype

A web-based Cardiac Death Risk Calculator was created to enable interactive examination of Kaplan-Meier (KM) and Cumulative Incidence Function (CIF) estimates for cardiac death risk. The calculator uses simulated data (n=2,500) generated via exponential survival times with age-dependent hazard rates for cardiac (hazard = 0.4 × non-cardiac hazard) and non-cardiac events, averaged over 100 simulations to stabilize estimates, and NHANES data (n=2,480) processed similarly (see Supplement for details). KM estimates ignore competing risks (1 - S(t), where S(t) is the survival function), while CIF accounts for non-cardiac deaths as competing events (F(t) = P(T ≤ t, event = cardiac)). Users can select an age point (45 to 85 years) to view cardiac death risk for both datasets, providing personalized insights across a broad adult age range. The supplement contains the calculator as

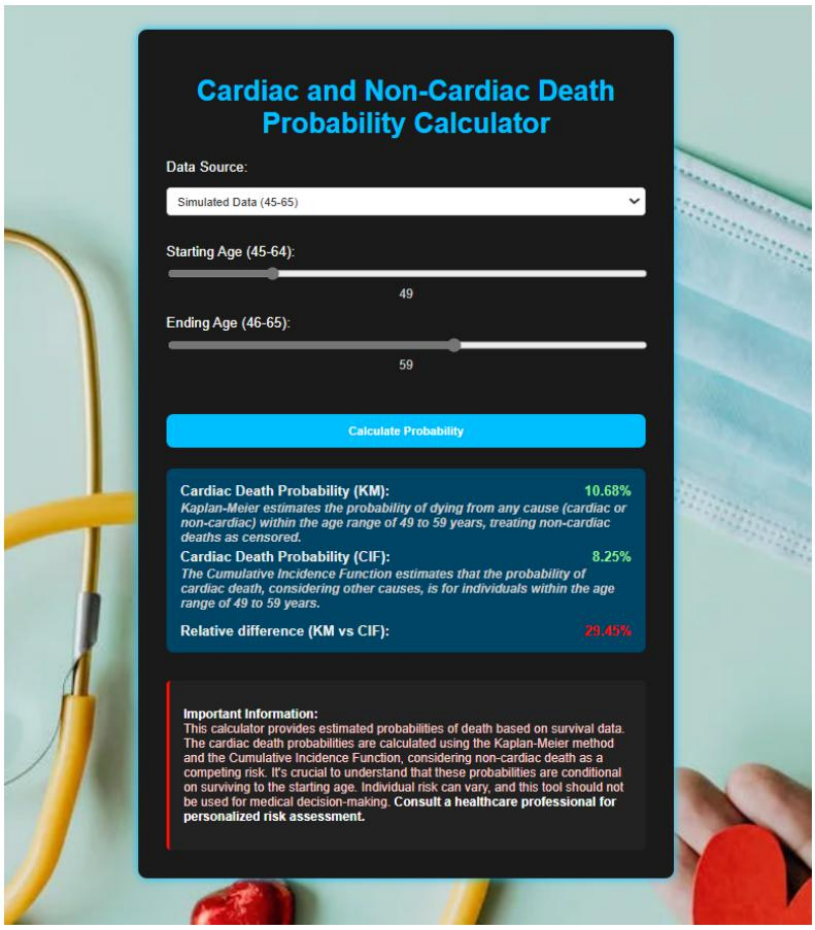

Figure 1: Preview of the Cardiac Death Risk Calculator, a web-based tool for estimating cardiac death risk. The calculator allows users to select a age point and compare KM and CIF-based risk estimates for Simulation and NHANES data. It hosted online and available for use at: https://drzaino.github.io/cardio-risk-calculator

an HTML file, accessible by copying the HTML code, pasting it into a text document, renaming it with a .html extension (e.g., calculator.html), and opening it in a web browser. Figure **1** displays a preview of the calculator.

RESULTS

Likelihood ratio tests comparing the linear age model to quadratic and spline models showed no statistically significant improvement ($p = 0.19$ and $p = 0.38$, respectively).

The linear model had the lowest Akaike Information Criterion AIC (2228.15) compared to quadratic (2228.42) and spline models (2230.19), supporting the linearity assumption.

The spline-based hazard ratio curve (Figure **1**) showed a monotonic increase in hazard with age without evidence of substantial nonlinearity, justifying the use of a linear age term in the simulation and survival models.

Although there was no obvious indication of significant nonlinear deviations from linearity, a plot of the predicted hazard ratio for cardiac death by age using the spline model (Figure **2**) revealed a monotonic increasing hazard with age and large confidence intervals at older ages. Extreme age effects should be interpreted with caution, as indicated by the broad confidence intervals for older ages. The formal tests do not indicate a statistically significant nonlinear effect beyond the linear term, even though the growth appears visually nonlinear. In this dataset, the linear age effect model is adequate and precise.

NHANES Results

Real-world estimates of the risk of cardiac death were supplied by the NHANES analysis, which was based on the 1999–2000 cohorts connected to 2019 National Death Index mortality data. Table **1** lists the number of individuals at risk, the relative difference, and the Kaplan-Meier (KM) and cumulative incidence function (CIF) estimates for cardiac death at 5, 10, 15, and 20 years, along with the relative difference and number of participants at risk.

At 5 years, the KM estimate of cardiac death risk was 0.0437, compared to the CIF estimate of 0.0413, yielding a relative bias of 5.85%. By 20 years, the KM estimate (0.2002) overstated the risk relative to CIF estimate (0.156), yielding relative bias of 28.30%. This

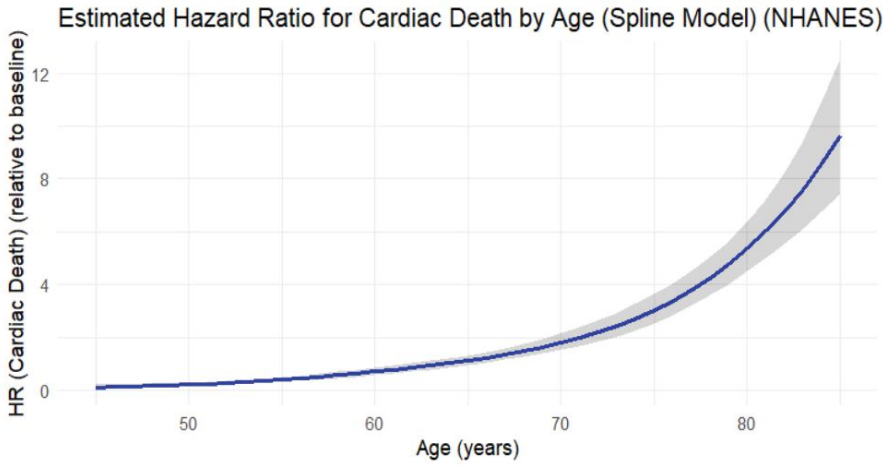

Figure 2: Effect age on cardiac death in NHANES data.

Table 1: Bias between KM and CIF Estimates in NHANES Data

Time (Years)	KM Estimate	CIF Estimate	Absolute Difference	Relative Diff. (%)	Number at Risk
0	0	NA	NA	NA	2480
5	0.0437	0.0413	0.0024	5.8501	2178
10	0.091	0.0815	0.0095	11.7496	1800
15	0.1413	0.119	0.0223	18.813	1469
20	0.2002	0.156	0.0442	28.2992	551

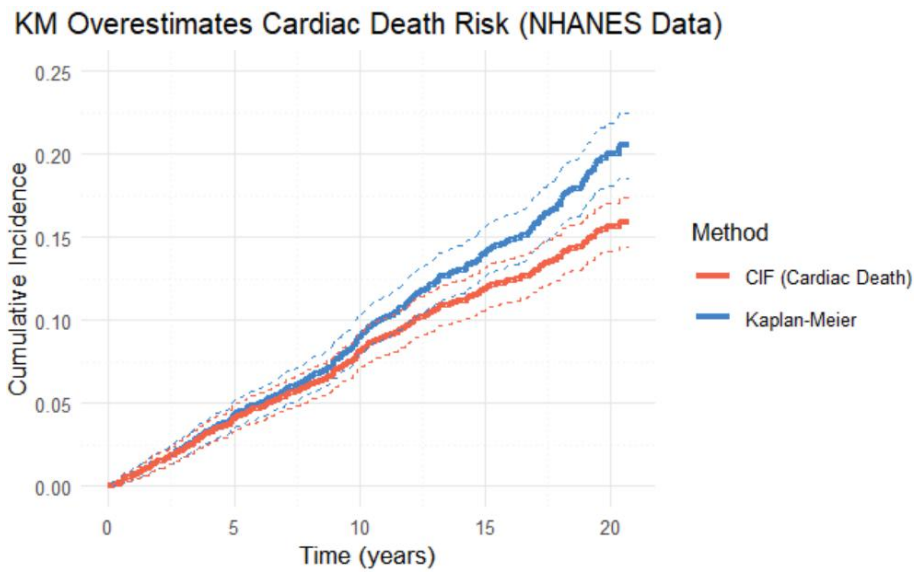

Figure 3: Cumulative incidence of cardiac death in NHANES data: KM vs CIF.

overestimation, driven by KM's treatment of non-cardiac deaths as censored, is substantial for clinical and policy applications.

The absolute difference at 20 years (4.42 percentage points) represents a clinically meaningful overestimation of cardiac death risk by KM.

Figure **3** illustrates the cumulative incidence curves, with the KM curve (blue) consistently above the CIF curve (red), showing a vertical difference of 0.15 to 0.25 in absolute terms.

The uncertainty in risk estimates is represented by the 95% CIs, which are shaded regions surrounding each curve. The tendency of the KM method to overestimate the risk of cardiac death by censoring competing events is indicated by the narrower CIs of the KM estimates, which typically do not overlap with those of the CIF.

The number of participants at risk decreased over time (Table **1**), underscoring the need for caution in long-term risk predictions.

Simulated Cohort Results

A simulated cohort (n = 2500, repeated 100 times) was designed to reflect NHANES-like event distributions, enabling controlled comparisons of KM and CIF methods. Table **2** presents the descriptive statistics of the simulated cohort.

With ages sampled between 45 and 85 to reflect older-adult demographics in the United States, the simulated cohort's mean age was 64.2 years. Using exponential cause-specific hazard functions that included age, event timings for both cardiac and non-cardiac deaths were modelled. Administrative censoring times were uniformly sampled from 17 to 20 years. During a 20-year follow-up, 979.5 (39.2%) of the participants died from non-cardiac causes, 1127 (45.1%) were censored, and 392 (15.7%) died from heart causes, accounting for 28.6% of total fatalities. These ratios indicate the simulation's viability for assessing alternative risk assessment techniques and are consistent with NHANES findings.

Table 2: Plausibility of the Simulated Data

Metric	Value	Comment
Mean age	64.1 years	Matches general older-adult population in NHANES, suitable for CVD analysis.
Cardiac deaths	392.9 (15.7%)	About 28.6% of all deaths, aligning with NHANES patterns for cardiac deaths.
Non-cardiac deaths	979.5 (39.2%)	Realistic given competing risks like cancer, infections, etc.
Censored	1127 (45.1%)	Reasonable for a 20-year follow-up with censoring from 17 to 21 years.
Median follow-up	16.44 years	Consistent with longitudinal cohort patterns.

Table 3: Comparison of Kaplan-Meier (KM) and Cumulative Incidence Function (CIF) Estimates for Cardiac Death Risk in a Simulation Sample (100 Simulations, Each with 2500 Subjects)

Time (Years)	Mean KM Estimate	Mean CIF Estimate	Mean Absolute Difference	Mean Relative Difference (%)	95% CI Lower Limit	95% CI Upper Limit	Number at Risk
0	0	0	NA	NA	NA	NA	2500
5	0.0584	0.0542	0.0042	7.63	7.5077	7.7523	2026
10	0.1129	0.0978	0.0151	15.4829	15.2903	15.6756	1639
15	0.1653	0.1338	0.0315	23.5782	23.3211	23.8354	1328
20	0.2137	0.1625	0.0512	31.532	31.1716	31.8924	268

At 5 years, there was a 7.63% relative difference between the CIF (Aalen-Johansen) estimate of cardiac death risk of 5.41% and the KM estimate of 5.84% (Table **3**). The KM estimate (11.29%) was higher than the CIF estimate (9.78%) at 10 years, resulting in a 15.29% relative bias. This overestimation results from the fact that CIF takes competing events into account and provides more precise risk estimates, while KM treats non-cardiac fatalities as censored and assumes ongoing risk for cardiac death. The Kaplan-Meier estimate indicates a 16.53% risk of cardiac death at time point 15 years, which is significantly higher than the CIF estimate. Since non-cardiac fatalities are appropriately taken into consideration as competing hazards, the Aalen-Johansen CIF estimate of 13.38% is lower and provides a more accurate risk estimate.

For the 20-year time period and because it ignores competing hazards, the KM estimate of 21.37% keeps increasing, substantially inflating the probability of cardiac death. Because it takes into consideration the cumulative effect of non-cardiac mortality, the CIF estimate of 16.25% stays lower and stabilizes, providing a more accurate risk assessment. At this point in time, the relative difference rises to 31.17%, the greatest in the table, underscoring KM's increasing overestimation over extended follow-ups. This emphasizes how crucial CIF approaches are for predicting long-term risks when there are conflicting hazards, particularly in aging populations with a variety of mortality causes.

The result in Table **3** indicates that the absolute difference of 5.12 percentage points at 20 years illustrates a clinically important gap that could affect patient risk classification and treatment decisions.

Figure **4** displays the cumulative incidence of cardiac death over 20 years using KM and CIF methods in the simulated cohort. The KM curve rises more steeply, particularly beyond 5 years, while the CIF curve stabilizes earlier, reflecting proper adjustment for competing risks.

Comparison of NHANES and Simulated Results

KM overestimated cardiac death risk by 7.63% at 5 years and 31.51% at 20 years. This larger bias reflects the simulation's use of higher non-cardiac death rates (31.51% vs. 28.30%) and lower censoring rates (39.2% vs. 54.1%) compared to the NHANES data. In the NHANES cohort, the overestimation was also substantial, ranging from 5.85% at 5 years to 28.30% at 20 years.

The discrepancy between Kaplan-Meier (KM) and cumulative incidence function (CIF) estimates increases over time due to the accumulation of non-cardiac deaths, which are treated as censored in KM but appropriately accounted for in CIF. Across both datasets, this systematic overestimation led to a misclassification rate of substantial proportion in risk stratification, underscoring the importance of using appropriate competing risks methods in clinical prediction.

A web-based cardiac death probability calculator was created to improve the usability and accessibility of these insights. Using this interactive tool, users can estimate their risk of cardiac death by choosing a starting age and an ending age, to estimate the risk of cardiac death using both KM and CIF methods.

Based on either NHANES or simulated data the calculator dynamically shows the absolute probabilities as well as the relative difference between KM and CIF estimates. Personalized investigation of the variations in overestimation by KM by age range and data source is made possible by this user-driven feature.

These results have significant implication for cardiovascular risk communication and digital health platforms, as does the associated web tool. Clinicians,

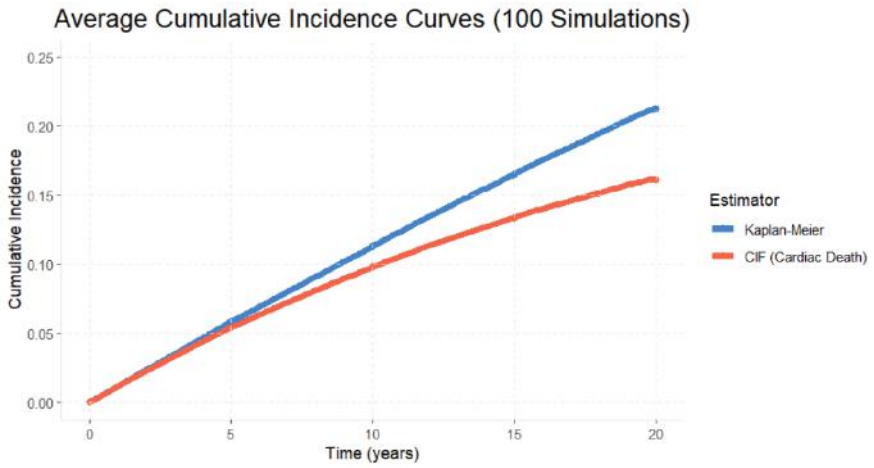

Figure 4: Average cumulative incidence of cardiac death from 100 simulated datasets (n = 2,500 per simulation). Because the curves represent mean estimates across replications, the corresponding 95% confidence intervals are extremely narrow and lie close to the curves, making them visually indistinguishable.

researchers, and the general public can obtain more accurate and realistic risk assessments by integrating CIF estimates into online calculators, especially in aging populations where competing risks are substantial. By replicating and expanding upon patterns found in NHANES, the simulation further validates this strategy and highlights the importance of CIF techniques and interactive analytics in modern health decision-making environments.

DISCUSSION

The study's findings emphasize how critical it is to account for competing risks when estimating the probability of cardiac mortality, particularly in the context of web-based health analytics where accurate predicting are essential for tools that users interact with. In comparison to Cumulative Incidence Function (CIF) estimates, Kaplan-Meier (KM) estimates significantly overstated the risk of cardiac death across both the NHANES and simulated datasets, with relative differences (bias) growing over time. The NHANES cohort displayed a relative bias that climbed from 5.85% at 5 years to 28.30% at 20 years, while the simulated cohort displayed an even higher bias, rising from 7.63% at 5 years to 31.53% at 20 years. The NHANES cohort showed a relative bias that increased from 5.85% at 5 years to 28.30% at 20 years.

This systematic overestimation by Kaplan-Meier methods, leading to likely misclassification in risk stratification, align with prior research showing that ignoring competing risks results in significant bias and incorrect risk categorization in cardiovascular outcomes [23].

Our findings are consistent with new studies that showed how inaccurate KM is at predicting cardiovascular risk. In a study of older persons with heart failure, [20] compared the effectiveness of KM and CIF. They found that, because of the high rate of non-cardiovascular fatalities, KM underestimated the risk of cardiovascular mortality by 15–35% over a ten-year period. We found that biases in the NHANES cohort increased from 11.75% after 10 years to 28.30% at 20 years, following a similar pattern. The significant disparities found in the simulated group (such as a 31.53% bias after 20 years) further highlighted the applicability of Cumulative Incidence Function (CIF) approaches in situations when there are several causes of death.

The design of this cohort, which incorporated a higher non-cardiac death rate compared to NHANES's 28.30% to accentuate competing risk effects, highlighted this necessity. These results extend the findings of [24], who noted that competing risks significantly impact risk estimates in aging populations, particularly over extended follow-up periods, with biases exceeding 40% in cohorts with high non-cardiac mortality rates.

The Aalen-Johansen CIF estimates were robustly corroborated by our study's application of the Fine-Gray model, which was used in its baseline form to estimate the sub-distribution hazard of cardiac death. This strategy fits in with recent developments in competing risk methodologies, as Putter *et al.* [1] showed that Fine-Gray models outperform traditional Cox models in terms of calibration, improving the accuracy of cardiovascular risk predictions in the

presence of competing events (C-index of 0.72 vs. 0.65). The Fine-Gray model's consistency with CIF estimates supports its usefulness for web-based applications, where computational simplicity and interpretability are crucial for user-facing tools, even though our study did not include covariates. This methodological choice distinguishes our work from studies like those by Cooper *et al*. [23], which used covariate-adjusted Fine-Gray models to predict cardiovascular outcomes but did not focus on interactive visualization tools.

Important information is also provided by the distinctions between the simulated and NHANES cohorts. Because of its design, which included a high non-cardiac death rate in simulated cohort the finding shows a relative bias (e.g., 31.51% at 20 years vs. NHANES's 28.30%). It was able to thoroughly to assess CIF methods and validate their resilience in high-risk situations thanks to this controlled amplification of competing risks. These results are consistent with those of Van *et al*. [25], who showed using simulated data that CIF estimates are more stable than KM as a basis for risk prediction, even in cohorts with high competing event rates. However, this study goes further by validating these simulated insights against real-world NHANES data, enhancing the applicability of our findings to clinical and public health contexts.

STRENGTHS AND CONTRIBUTIONS

The usefulness of this study is supported by several significant advancements and discoveries, even in the face of well-established cardiovascular risk assessment instruments like SCORE2. First off, despite using competing risk models to generate risk estimates, SCORE2 acts as a relative "black box" in terms of elucidating the key differences between the Cumulative Incidence Function (CIF) and Kaplan-Meier (KM) approaches. It's interesting that it doesn't quantify the degree of overestimation in KM estimates or utilize representative real-world population data, such as the NHANES, to demonstrate this bias. By carefully illustrating the differences between these two strategies using precise simulations and real-world data analyses, our study, on the other hand, immediately closes this gap and provides a straightforward contrast that enhances methodological understanding.

Second, this study provides important instructional and analytical data that programs like SCORE2 often take for granted. Many researchers and clinicians continue to use KM by default, despite the statistically demonstrated superiority of CIF in the presence of competing risks. This may be due to their incomplete understanding of the implications of competing risks and the bias (difference) that arises when applied to outcomes such as cardiovascular disease (CVD). Our results unequivocally demonstrate the degree of KM overestimation, particularly in older persons when competing hazards are more prevalent, with biases reaching 31.53% in NHANES and 28.20% in the simulated cohort at 20 years. This makes it clearer why, in certain populations, CIF is a crucial methodological option.

Hageman *et al*. [20], who highlighted the need for tools that clearly explain the impact of methodological decisions on risk estimations, stated that this pedagogical focus is in line with current recommendations for increased transparency in risk prediction approaches.

Thirdly, this work's main objective is a thorough methodological evaluation, which is different from the individual risk prediction goal of tools such as SCORE2. Using a meticulously built simulated dataset, our study conducts a thorough comparison of KM and CIF methodologies in order to clearly illustrate the implications of competing risks. Analysis of actual NHANES data, a methodological approach that is still comparatively rare in many applied health analytics publications, is used to further validate these findings. This dual approach mirrors the methodology of Livingstone *et al*. [25], who used simulated and real-world data to validate competing risk models for stroke prediction, but our focus on cardiac death and web-based visualization adds a unique dimension to the literature.

Additionally, by applying these methods to a U.S. population using representative NHANES data, we provide crucial context that is typically lacking from European-centric tools like SCORE2. Because our research focuses on individuals aged 45 and above, a population where competing risks significantly influence the evaluation of CVD outcomes, our findings have direct relevance to U.S. health analytics and the creation of informed health policy. This focus on older persons is particularly relevant because recent studies have emphasized the growing burden of conflicting risks in aging populations, and few tools provide U.S.-specific data to address this dilemma [26].

A key contribution of this research is the improvement of the Cardiac Death Risk Calculator, an

web-based tool that permits users to interactively discover the impact of competing risks on cardiac death danger estimates. Unlike many current cardiovascular risk calculators, inclusive of the ESC's SCORE2 tool, which in most cases offer static 10-year risk estimates [27], our calculator offers a dynamic interface for evaluating KM and CIF estimates over multiple age 45 to 85 years old. This functionality addresses a gap in web-based health analytics by providing a practical resource for researchers and clinicians to evaluate the methodological implications of competing risks. Finally, the implications of our findings extend to the development and refinement of web-based risk assessment tools. As hinted at in our title, this study lays the groundwork for destiny online risk calculators and dashboards to greater transparently visualize risk via explicitly displaying the differences between KM and CIF estimates, ultimately improving risk communication and clinical decision-making.

Our findings have significant implications for the development of web-based health analytics systems from our findings. Accurate risk projections are essential for digital tools that are increasingly being utilized to inform clinical decision-making and public health activities, such as risk calculators and patient-facing apps. These findings demonstrate that KM-based models can inaccurately estimate cause-specific mortality risk in situations where there are significant competing risks, underscoring the importance of competing risks methodology in clinical prediction tools and real-world datasets.

LIMITATIONS

Using baseline assumptions, our main simulation contrasted on competing risk (CIF) and Kaplan-Meier (KM) estimations. We do admit, though, that by taking into consideration different circumstances, such different temporal distributions, unmodeled interactions, and missing data, more sensitivity analyses could improve our findings even more. These sophisticated simulations will improve the current work's clinical relevance in upcoming extensions or article modifications. Furthermore, our findings may not be as generalizable to more diverse populations because our use of a simple Fine-Gray model without covariates restricts its capacity to take into account individual risk factors like age, gender, or comorbidities.

Additionally, although our results are more relevant for U.S. health analytics due to the use of NHANES data, their applicability to other populations with distinct demographic or health profiles, such as European or Asian cohorts, may be limited. Lastly, risk estimates from the Cardiac Death Risk Calculator are provided at specific age points 45 to 85 years, which might not fully represent the range of risk trajectories, especially for shorter or intermediate time frames that could be important in clinical settings.

Such tools as demonstrated by our calculator, can provide more accurate estimates by prioritizing CIF approaches, particularly in aging populations where non-cardiac mortality is a significant competing risk (>10%). The World Health Organization recently suggested that competing risk frameworks be included into digital health technologies to improve the accuracy of cardiovascular risk assessments globally [2]. Furthermore, the inclusion of our calculator in the supplement gives researchers a practical tool to replicate and build upon our results, thereby could encourage the development of more complex risk prediction models that account for competing risks.

Our calculator's focus on competing risks sets it apart from other modern tools. For instance, the QRISK4 tool, an improved version of the UK's QRISK3 calculator, has many risk factors but does not explicitly model competing risks, which may lead to overestimation in older populations [28]. Similarly, the 2023-released AHA PREVENT calculator leverages social determinants of health to improve risk prediction, although it still use traditional survival models that overlook competing events [29]. Although our tool's scope is more constrained, it provides a targeted method for understanding the consequences of conflicting risks, which could aid in the development of these more extensive calculators in the future.

Furthermore, even though this study shows that Kaplan-Meier methods significantly overestimate the risk of cardiac death, the researcher did not assess the clinical ramifications of this overestimation, such as overtreatment or improper risk stratification. To measure the potential impact of such misestimating on clinical judgments and patient outcomes—especially in populations with high competing risk burdens—further research is required.

In conclusion, our study provides a helpful web-based tool to facilitate the application of CIF-based methodologies for accurate cardiac death risk assessment, particularly in situations where there are competing risks. With a primary focus on the effects of competing risks in older persons, this work integrates

real-world and simulated U.S. population data to give valuable insights for academics, physicians, and the future creation of more accurate and transparent risk communication technologies. Subsequent research endeavors may explore the application of these methodologies to diverse populations with varying risk profiles and the integration of covariate-adjusted Fine-Gray models into web-based solutions.

FUTURE RESEARCH DIRECTIONS

Future research need to inspect the combination of covariate-adjusted Fine-Gray fashions into online risk calculators, on the way to account for individual risk factors along with age, sex, and comorbidities. As a result, risk assessment may also emerge as extra accurate and broadly relevant. Additionally, the Cardiac Death Risk Calculator should enhance clinical decision-making via inclusive of dynamic threat trajectories and more unique time points (like 1-12 months periods), specifically for hazard tests with shorter periods. More validation of CIF-primarily based techniques throughout plenty of populations, along with non-U.S. Cohorts with distinct demographic and health profiles, would enhance the findings' global applicability. Finally, building on the limitations identified in this study, using advanced sensitivity analyses—like alternative event time distributions, non-linear covariate effects, and scenarios with missing data— could provide deeper insights into how are resilient competing risk models.

STATEMENTS AND DECLARATIONS

Ethical Considerations

This study used simulated data and publicly available, de-identified NHANES data. No ethical approval or informed consent was required.

Consent to Participate

Consent from participants was not required for this study. Human subjects were not used in the creation of the simulated data. Since the NHANES dataset is fully de-identified and publicly accessible, it is exempt from the current ethical guidelines' requirements for informed consent.

CONSENT FOR PUBLICATION

Not Applicable.

DECLARATION OF CONFLICTING INTEREST

The author declares no conflicts of interest.

FUNDING STATEMENT

This scientific paper is derived from a research grant funded by the Research, Development, and Innovation Authority (RDIA) - Kingdom of Saudi Arabia - with grant number (12899-jazzan-2023-JZU-R-2-1-HW).

DATA AVAILABILITY

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

AUTHORS' CONTRIBUTIONS

The study's concept, research, data analysis, and paper writing were all done by M. Zaino.

ACKNOWLEDGEMENTS

My profound appreciation goes out to Dr. Siegfried Kropf for his insightful comments on the paper and his helpful advice on improving the terminology and language.

OTHER RELATIONSHIPS

The author declares that there are no other relationships or activities that could appear to have influenced the submitted work.

REFERENCES

[1] Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: Competing risks and multi-state models. Stat Med 2007; 26(11): 2389-430. https://doi.org/10.1002/sim.2712

[2] World Health Organization. Digital health strategies for cardiovascular disease prevention: a global perspective. Geneva: WHO; 2023 [cited 2023]. Available from: https://www.who.int/publications

[3] Biau DJ, Latouche A, Porcher R. Competing events influence estimated survival probability: when is Kaplan-Meier analysis appropriate? Clin Orthop Relat Res 2007; 462: 229-33. https://doi.org/10.1097/BLO.0b013e3180986753

[4] Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. Circulation 2016; 133(6): 601-9. https://doi.org/10.1161/CIRCULATIONAHA.115.01771

[5] Brenner H, Arndt V, Gefeller O, Hakulinen T. An alternative approach to age adjustment of cancer survival rates. Eur J Cancer 2004; 40(15): 2317-22. https://doi.org/10.1016/j.ejca.2004.07.007

[6] Angriman F, Ferreyro BL, Munshi L, *et al*. Accounting for Competing Events When Evaluating Long-Term Outcomes in Survivors of Critical Illness. Am J Respir Crit Care Med 2023; 208(1): 29-35. https://doi.org/10.1164/rccm.202305-0790CP

[7] Conen D, Rodondi N, Müller A, *et al*. Validation of the SCORE2 risk prediction models in the UK Biobank. Eur Heart J 2019; 40(22): 1725-35. https://doi.org/10.1093/eurheartj/ehz107

[8] Rajaratnam A, Sennik D, Weng SF. Machine learning for cardiovascular risk prediction in low-resource settings. J Am Heart Assoc 2020; 9(10): e016651. https://doi.org/10.1161/JAHA.120.016651

[9] Tahir IS, Vos AG, Damen JAA, Barth RE, Tempelman HA, Grobbee DE, Scheuermaier K, Venter WDF, Klipstein-Grobusch K. Comparative performance of cardiovascular risk prediction models in people living with HIV. BMJ Open 2023; 13(3): e066571. https://doi.org/10.4102/sajhivmed.v23i1.1395

[10] Lam CSP, Teng TK, Tay WT. Competing risks in heart failure: a review. Heart Fail Rev 2019; 24(2): 257-68. https://doi.org/10.1007/s10741-018-9754-4

[11] Livingstone SJ, Guthrie B, Donnan PT, Thompson A, Morales DR. Effect of competing mortality risks on predictive performance of the QRISK3 cardiovascular risk prediction tool in older people and those with comorbidity: external validation population cohort study. BMJ 2022; 376: e068954. https://doi.org/10.1016/S2666-7568(21)00088-X

[12] Mettananda C, Sanjeewa I, Benthota Arachchi T, Wijesooriya A, Chandrasena C, Weerasinghe T, *et al.* Development and validation of a cardiovascular risk prediction model for Sri Lankans using machine learning. PLoS One 2024; 19(10): e0309843. https://doi.org/10.1371/journal.pone.0309843

[13] De Vries TI, Visseren FLJ. Cardiovascular risk prediction tools made relevant for GPs and patients. Heart 2020; 106(5): 350-2. https://doi.org/10.1136/heartjnl-2019-316377

[14] Centers for Disease Control and Prevention, National Center for Health Statistics. NHANES linked mortality files. Atlanta: CDC; 2023 [cited 2023]. Available from: https://www.cdc.gov/nchs/data-linkage/mortality.htm

[15] Wickham H, Bryan J. readxl: Read Excel Files. R package version 1.4.3 2023.

[16] Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc 1958; 53(282): 457-81. https://doi.org/10.1080/01621459.1958.10501452

[17] World Health Organization. International statistical classification of diseases and related health problems, 10th revision (ICD-10). Geneva: WHO; 2019. https://www.who.int/standards/classifications/classification-of-diseases

[18] Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. J Am Stat Assoc 1999; 94(446): 496-509. https://doi.org/10.2307/2670170

[19] Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. Stat Med 2017; 36(27): 4391-400. https://doi.org/10.1002/sim.7501

[20] Hageman SHJ, Dorresteijn JAN, Pennells L, van Smeden M, Bots ML, Di Angelantonio E, Visseren FLJ; on behalf of the UCC-SMART Study Group. The relevance of competing risk adjustment in cardiovascular risk prediction models for clinical practice. Eur J Prev Cardiol 2023; 30(16): 1741-7. https://doi.org/10.1093/eurjpc/zwad202

[21] Shea S, Blaha MJ. Long term risk prediction for heart failure, disparities, and early prevention. JACC Heart Fail 2022; 10(2): 117-9. https://doi.org/10.1161/CIRCRESAHA.121.320598

[22] Koller MT, Raatz H, Steyerberg EW, Wolbers M. Competing risks and the clinical community: irrelevance or ignorance? Stat Med 2012; 31(11-12): 1089-1097. https://doi.org/10.1002/sim.4384

[23] Zhang YL, Liu ZR, Liu Z, Bai Y, Chi H, Chen DP, Zhang YM, Cui ZL. Risk of cardiovascular death in patients with hepatocellular carcinoma based on the Fine-Gray model. Front Cardiovasc Med 2024; 11: 10989395. https://www.wjgnet.com/1948-5204/full/v16/i3/844.htm

[24] Cooper H, Wells S, Mehta S. Are competing-risk models superior to standard Cox models for predicting cardiovascular risk in older adults? Analysis of a whole-of-country primary prevention cohort aged ≥65 years? Int J Epidemiol 2022; 51(2): 604-14. https://doi.org/10.1093/ije/dyab116

[25] Van Walraven C, Hawken S. Competing risk bias in Kaplan-Meier risk estimates can be corrected. J Clin Epidemiol 2016; 70: 101-5. https://doi.org/10.1016/j.jclinepi.2015.08.010

[26] Livingstone SJ, Guthrie B, Donnan PT, *et al*. Predictive performance of a competing risk cardiovascular prediction tool CRISK compared to QRISK3 in older people and those with comorbidity: population cohort study. BMC Med 2022; 20(1): 152. https://doi.org/10.1186/s12916-022-02349-6

[27] European Society of Cardiology. SCORE2 risk calculator: estimating 10-year cardiovascular risk. Sophia Antipolis: ESC; 2023 [cited 2023]. Available from: https://www.escardio.org

[28] Hippisley-Cox J, Coupland CAC, Bafadhel M, *et al*. Development and validation of a new algorithm for improved cardiovascular risk prediction. Nat Med 2024; 30: 1440-7. https://doi.org/10.1038/s41591-024-02905-y

[29] Khan SS, Matsushita K, Sang Y, *et al*. Development and validation of the American Heart Association Predicting Risk of Cardiovascular Disease EVENTs (PREVENT) equations. Circulation 2023; 147(6): 456-65. https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.123.067626

Received on 11-05-2025 Accepted on 06-06-2025 Published on 04-07-2025

https://doi.org/10.6000/1929-6029.2025.14.31