Clinical Epidemiology

Dovepress
Taylor & Francis Group

Open Access Full Text Article

REVIEW

An Introduction to Competing Risks in Epidemiology

Henrik Toft Sørensen[1,2], Erzsébet Horváth-Puhó[1], Janet L Peacock[1,2]

[1]Department of Clinical Epidemiology, Center for Population Medicine, Aarhus University Hospital and Aarhus University, Aarhus, Denmark; [2]Department of Epidemiology, Geisel School of Medicine at Dartmouth, Dartmouth College, Hanover, NH, USA

Correspondence: Henrik Toft Sørensen, Department of Clinical Epidemiology, Center for Population Medicine, Aarhus University Hospital and Aarhus University, Olof Palmes Allé 43-45, Aarhus N, DK-8200, Denmark, Email hts@clin.au.dk

Abstract: Competing risks arise when people are at risk of multiple mutually exclusive events, such that the occurrence of one event alters the probability of others. In research, ignoring competing risks can lead to biased estimates. We outline key approaches for analyzing competing risk data, focusing on their assumptions, interpretations, and epidemiological and clinical relevance. The Aalen-Johansen estimator, a non-parametric method for estimating the cumulative incidence function, provides an alternative to the naïve Kaplan-Meier estimator when competing events are present. The cause-specific hazard model estimates the instantaneous risk of a specific event type, treating competing events as censored, and is used for etiologic research. The Fine-Gray subdistribution hazard model directly models the cumulative incidence function, thus offering a clinically interpretable measure of absolute risk. We also discuss the use of composite endpoints, which combine several event types to increase statistical power, and highlight their limitations in clinical interpretation. By comparing these methods and illustrating their applications through analyses of the association between venous thromboembolism and arterial events, this review aims to guide researchers, particularly junior researchers, in selecting appropriate strategies for valid and meaningful analysis of competing risks in clinical and epidemiological studies.

Keywords: competing risks, epidemiology, Aalen-Johansen estimator, cause-specific hazard model, fine-gray subdistribution hazard model

Introduction

In epidemiology and clinical research, survival analyses are commonly used in time-to-event analyses.[1,2] "Competing risks" refers to situations in which multiple potential events can occur, but the occurrence of one event prevents the occurrence of the others.[3] A competing event is the actual occurrence of one of those competing risks. For example, a patient who dies of acute myocardial infarction cannot experience a stroke thereafter. In this context, myocardial infarction can be considered a competing risk. If a patient experiences myocardial infarction, it might alter the risk profile or lead to death before a stroke can occur. Competing risks can occur in both randomized and observational studies. "Semi-competing risks" refers to situations in which two types of events can occur, but one event might preclude the occurrence of the other, whereas the reverse is not possible.[4]

The aim of the paper is to introduce clinicians and researchers to the basic concepts of competing risks in epidemiology. Kaplan-Meier curves and hazard ratios are often misapplied and misinterpreted when competing risks are present, and much of the statistical literature emphasizes advanced methods without connecting them to basic epidemiological and clinical principles. With the ageing population and increasing multimorbidity, competing risks are common in population and clinical epidemiology.[5] Failure to use appropriate methods can lead to severely biased estimates, with implications for clinical guidelines and patient care. For example, applying Fine-Gray models or cause-specific hazard models without understanding how these measures relate to core epidemiological concepts can be problematic. In this paper, we present key epidemiological principles relevant to competing risk analysis, illustrated through a common example based on Danish registry data.

https://doi.org/10.2147/CLEP.S574928
Received: 17 October 2025
Accepted: 15 January 2026
Published: 2 February 2026

Basic Concepts

Before we discuss how to handle competing risks in statistical analyses, we will first cover several basic epidemiological concepts.

Epidemiological and clinical studies are often broadly divided into etiologic studies or prediction studies.[5,6] An etiologic study is used to identify and understand the causal relation between a risk factor and an outcome. The aim is to examine whether a specific exposure or factor causes a particular health outcome. In this type of research, we adjust for potential confounding factors to isolate the direct effect of the exposure on the outcome (Figure 1). A classic example is examining whether smoking causes lung cancer by comparing the incidence of lung cancer in a cohort of smokers versus a cohort of non-smokers, while controlling for other variables, such as age and sex.

The direct effect refers to the impact of an exposure on an outcome that is not mediated by other variables. A mediator is a variable that lies on the causal pathway between an exposure and an outcome, transmitting part or all of the exposure's effect on the outcome.[6] The indirect effect captures the proportion of the exposure's impact on the outcome that operates through one or more mediators (Figure 2). For example, the reduced risk of cardiovascular disease associated with high physical activity may be partly explained by lower blood pressure and body mass index, representing the indirect effect.

One aim of a prediction study is to develop models that can predict the likelihood of an outcome using multiple predictors.[6,7] These studies aim to forecast future events, and the goal is to develop a model that can accurately predict outcomes, regardless of whether the predictors are causally associated with the outcome. Hence in summary, etiologic studies focus on understanding the causes of health outcomes, whereas prediction studies focus on forecasting the likelihood of these outcomes according to various predictors (Figure 1).

Cohorts can be classified as either closed (fixed) or open (dynamic).[8] A closed cohort is defined at baseline when the membership of the cohort is fixed at the start of the study. After the cohort is defined, and follow-up begins, no additional participants can be added. Therefore, the number of participants might decrease because of events such as death, loss to follow-up, or development of the outcome being studied over time. In closed cohorts, only outcomes/events and censoring can occur. In contrast, in an open cohort, new members can enter the cohort at different times during the study period, and existing members can leave.[8] The cohort size can thus vary over time as participants enter and exit the study (Figure 3).

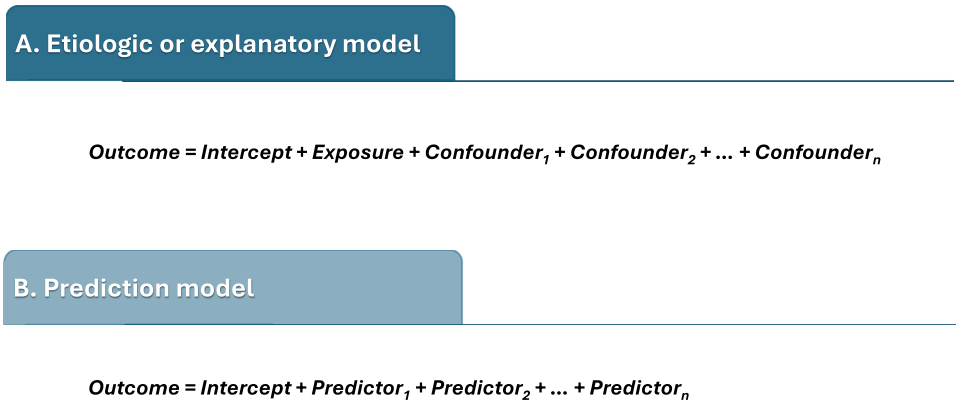

Figure 1 Description of the two basic epidemiological models (etiologic and prediction models).

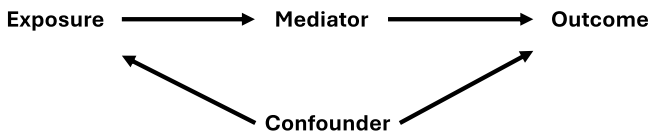

Figure 2 Diagram of confounding and mediation.

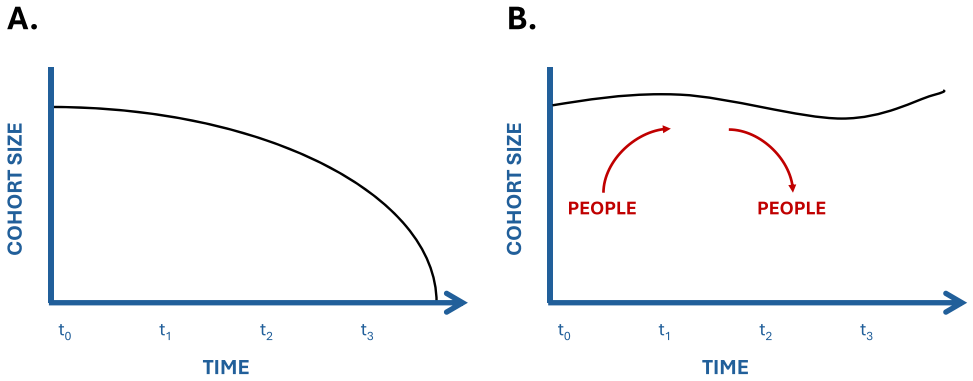

Figure 3 Illustration of closed (fixed) and open (dynamic) cohorts. (**A**) CLOSED (fixed) COHORT. (**B**) OPEN (dynamic) COHORT.

Censoring refers to a situation in which the event of interest (for example, death or disease) has not occurred in some participants at the end of the study, at the point at which the data are analysed, or some participants are lost to follow-up.[9] In these situations we do not know what happens to censored participants after the censoring point but survival methods can be used to combine their survival data up to the censoring point with survival data from participants who have experienced the event of interest, to estimate overall survival probabilities and survival curves. These types of studies are therefore modeling the length of time that a participant remains event-free and so are often known as time-to-event studies.

Censoring can be classified as right, left, or interval censoring. Right censoring occurs when follow-up ends or a participant is lost to follow-up before the event of interest occurs. Left censoring arises when the outcome occurs before the observation begins and the exact time of occurrence is unknown. Interval censoring occurs when the event happens between two observation points, but the precise timing is unknown (Figure 4).[10]

In a dynamic cohort, the incidence rate (hazard rate) can be calculated as the rate at which new cases of a disease occur in a population over a specified period. This rate can be measured as cases or outcomes per person-time. Risk or cumulative incidence is the probability that a person will develop the disease over a specific time period. This dimensionless measure ranges from 0 to 1. The incidence rate can be converted to risk if the follow-up time is consistent, and the population is stable (Table 1).

If the likelihood that a participant or observation is censored during follow-up does not depend on the reason for censoring, standard survival analysis methods such as Kaplan-Meier and traditional Cox proportional hazards regression analysis can be used. This censoring is called non-informative censoring.[11] In the presence of competing risk, however,

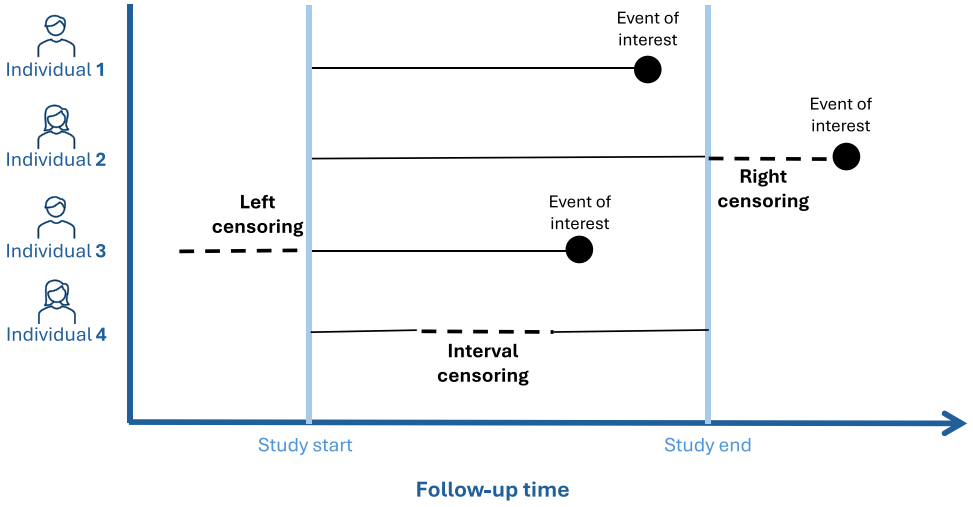

Figure 4 Graphical illustration of right, left, and interval censoring.

https://doi.org/10.2147/CLEP.S574928

Table 1 Association Between Risk and Incidence Rate

Risk	The probability of an outcome occurring in a specific time period
Incidence Rate	The number of new cases per unit of person time
Time	Duration over which the risk is being calculated
$\text{Risk} = 1 - e^{-(\text{Incidence Rate} \times \text{Time})}$	

the time to competing event and time to censoring might not be independent, because the exposure of interest might be associated with one or more competing risks. This type of censoring is called informative censoring. In such situations, use of Kaplan-Meier or traditional Cox regression might lead to biased estimates and hence invalid conclusions. Under these conditions, the Kaplan-Meier estimate of the cumulative incidence is interpretable as a cumulative incidence among participants who do not have a competing risk event (Figure 5A).

Intercurrent events are events that occur after the start of follow-up but before the outcome of interest and can influence either the exposure or the outcome.[12] Examples include treatment changes, development of another disease, or competing risks such as death from unrelated causes. An intercurrent event is therefore more broadly defined than a competing event, which specifically refers to events that prevent the outcome from occurring, eg death from another cause.[12]

In the presence of competing risks, the Kaplan-Meier method can lead to overestimation of the event probability because the Kaplan-Meier estimator treats competing risks as non-informative censored data. Consequently, if participants could be followed beyond the censoring date, the probability of the outcome is incorrectly assumed to be the same among those who are not censored at that time. Therefore, when a competing event occurs, it is not considered an event of interest; instead, the individual is considered no longer at risk of the primary event. For example, if we are studying time-to-death from cancer, and a patient dies from an acute myocardial infarction (a competing event), the Kaplan-Meier estimator would treat this patient as censored at the time of the acute myocardial infarction. The estimate would be biased because it ignores that the patient can no longer die from cancer after dying due to the acute myocardial infarction.

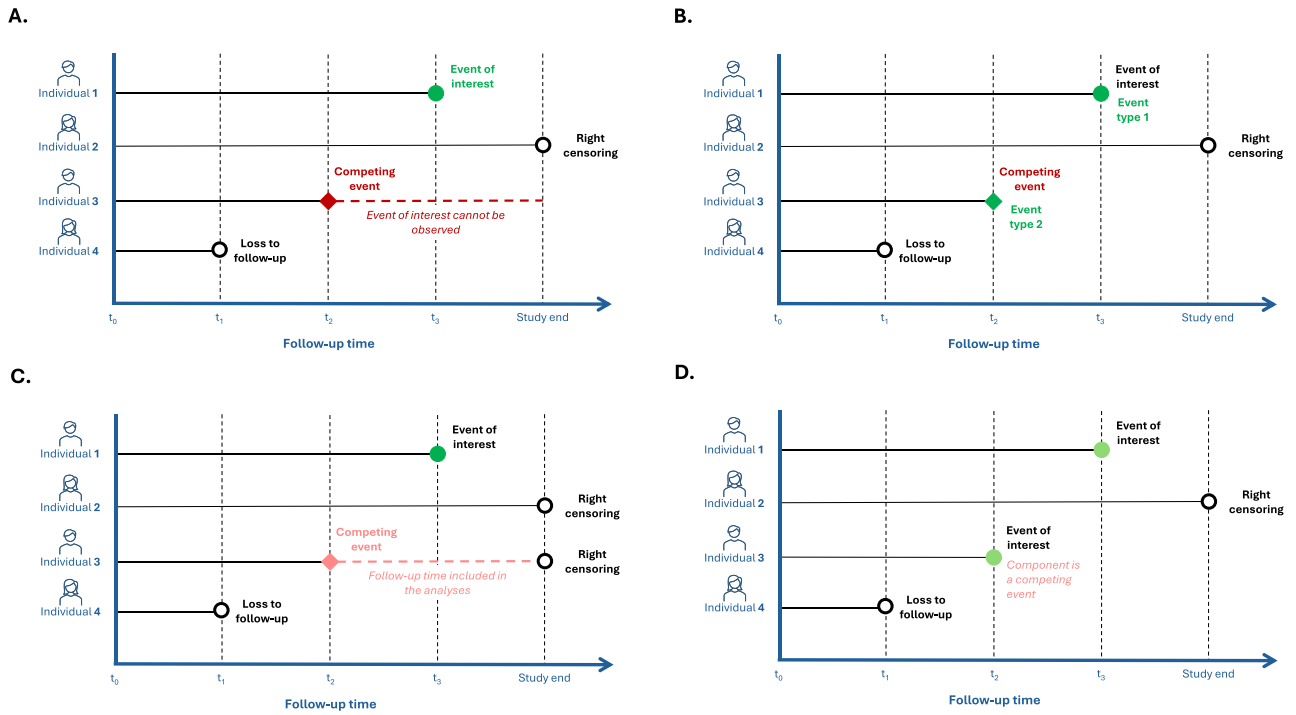

Figure 5 Competing risks: problem and solutions. (**A**) Illustration of the competing risks. (**B**) Aalen-Johansen cumulative incidence and cause-specific hazard model (**C**) Subdistribution hazard model. (**D**) Composite endpoint. *The green color denotes the event of interest, and the red color denotes the competing event.*

https://doi.org/10.2147/CLEP.S574928

In the context of competing risks, the association between rate and risk becomes more complex, and the direct 1-to-1 relation between the hazard rate and the cumulative risk is lost.[13,14] Consequently, how covariates affect the hazard might not directly translate to how they affect the cumulative incidence. In traditional Cox regression analysis, the logarithm of the relative hazard is being modeled, and the key assumptions are that the hazards are proportional over time, and that the censored participants have the same outcome rate as the non-censored participants. If this is not true, ie, the censoring is informative, the traditional Cox regression analysis will provide biased estimates. This issue is particularly relevant in the presence of competing risk and for any other reason the censoring is informative such as in a trial where patients drop out due to treatment side effects.

Example: Venous Thromboembolism and Arterial Cardiovascular Events

The association between venous thromboembolism (VTE) and arterial events is well established.[15] To illustrate the challenges of competing risk and describe potential solutions, we conducted a dynamic cohort analysis between January 1, 1996, and December 31, 2021, by using population-based health registries from Denmark.

We identified all patients with a first-time diagnosis of VTE in the Danish National Patient Registry, which has recorded all hospital discharge diagnoses in Denmark since 1977, and all outpatient clinic and emergency department visits since 1994. Using the Danish Civil Registration System, we randomly sampled a comparison cohort from the general Danish population with replacement,[16] matched on sex, year of birth, and year of VTE diagnosis, in a ratio as high as 5:1.

The primary outcome event of interest was a primary or secondary diagnosis of major adverse cardiovascular events (MACE; further described below), defined as a composite of myocardial infarction, ischemic stroke, or cardiovascular death. Deaths from non-cardiovascular causes were considered competing events. Causes of death were obtained from the Danish Cause of Death Register.[17]

The VTE cohort comprised 156,245 patients with a first-time VTE diagnosis between 1996 and 2021. The median age was 68 years (interquartile range: 55–78) at VTE diagnosis, and 52% were women. The comparison cohort included 468,740 people with comparable baseline characteristics. Patients with a VTE diagnosis and comparison individuals were followed from the date of VTE diagnosis or the index date until the first occurrence of MACE, non-cardiovascular death, emigration, loss to follow-up, December 31, 2021, or a maximum of 10 years of follow-up. During a median follow-up of 5.9 years (IQR: 2.4–10.0 years), 29,665 (19%) patients with VTE and 69,220 (15%) comparison cohort members were registered as having MACE. A total of 75,385 (48%) patients with VTE and 140,565 (30%) people in the comparison cohort died by non-cardiovascular causes during follow-up.

Naïve Kaplan-Meier curves of MACE were created, and observations were censored at the time of the competing event (non-cardiovascular death). The 10-year risk of MACE was 30.7% (95% confidence interval [CI]: 30.4–31.0%) in the VTE cohort compared with 23.2% (95% CI: 23.0–23.3%) in the matched comparison cohort (Figure 6A).

Approaches to Competing Risks in Survival Analysis

Aalen-Johansen Cumulative Incidence

The Aalen-Johansen approach is recommended for examining competing risks, because, in comparison to the Kaplan Meier approach, it provides a more valid estimation of the cumulative incidence of events in the presence of competing risks (Figure 5B).[13,14,18] The Aalen-Johansen approach is a non-parametric method that makes no assumptions about the underlying distribution of the times to event. This approach directly estimates the cumulative incidence function, which represents the probability of specific events occurring at a given time, considering the presence of competing events.[14] This estimator accounts for the ability of the competing event to preclude the occurrence of the event of interest, in contrast to the classic Nelson-Aalen approach, which is used for estimating the cumulative hazard in a single event study. The Aalen-Johansen approach avoids the overestimation of the event probability that can occur with methods such as the Kaplan-Meier estimator.

Using the example above, we applied the Aalen–Johansen cumulative incidence function estimator to calculate the cumulative incidence of MACE (event type 1), while treating non-cardiovascular death as a competing event (event type 2). The cumulative risk of MACE was 23.5% (95% CI: 23.3–23.8%) in the VTE cohort and 20.2% (95% CI: 20.1–20.3%) in

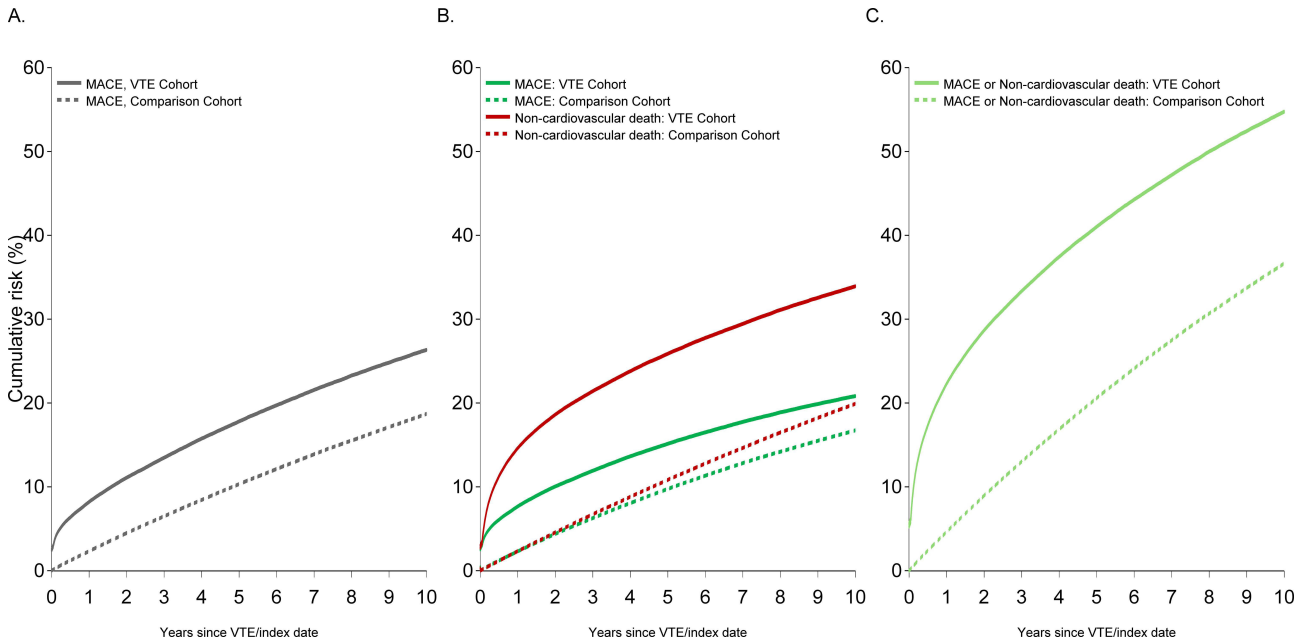

Figure 6 Illustrative example: venous thromboembolism and arterial cardiovascular events. (**A**) Naïve Kaplan-Meier Method. (**B**) Aalen-Johansen Cumulative Incidence. (**C**) Composite Endpoint.

the comparison cohort. Further, the method allows us to create cumulative incidence curves for the competing event (non-cardiovascular death, event type 2) (Figure 6B).

The two most used regression analyses for competing risks are the cause-specific hazard model and the Fine-Gray model.[19–24]

The Cause-Specific Hazard Model

The cause-specific hazard model should be used in the context of competing risk in etiologic research, because it enables differentiation among various types of events (eg, different causes of death),[19] whereas the binary Cox model used to measure disease rates and risks is based on the assumption of only one type of outcome. The cause-specific hazard model is a semi-parametric method that makes no assumptions about the distribution of times-to-events but assumes that the hazards are proportional. By modeling the hazard for each specific cause, researchers can understand how covariates affect the risk of each event separately (Figure 5B). The cause-specific hazard model, often based on the Cox proportional hazards model, allows the effects of covariates to differ by event type.[20] Thus, the effect of a covariate on one type of event can be different from its effect on another type of event. By focusing on cause-specific hazards, the analysis can provide insight into the mechanisms underlying each event type, but the model is based on an assumption that the competing risks are independent of each other and interpretation can therefore be difficult. The cause-specific hazard function provides the hazards for a specific cause (exposure) but does not provide the probability of an outcome. The cause-specific model is not always suitable for all study populations, particularly when the assumption of proportional hazards is not fulfilled. However, this method can help in understanding the pathways leading to different outcomes and thereby informing targeted interventions.

The Fine-Gray Subdistribution Hazard Model

The Fine-Gray subdistribution hazard model can be used in the presence of competing risks, to estimate the cumulative incidence of an event or predict a prognosis.[25] This is a semi-parametric method that makes no assumptions about the distribution of the times-to-events. Unlike traditional survival models (eg, Cox regression), which treat competing events as censored, the Fine-Gray model retains individuals who experience competing events in the risk set, adjusting the hazards accordingly (Figure 5C). The Fine-Gray model therefore provides a more accurate estimate of the cumulative

incidence function than methods such as Kaplan Meier. For example, when the aim is to predict the likelihood of an event, the Fine-Gray model allows the inclusion of covariates to predict the subdistribution hazards that directly relate to cumulative incidence. However, the Fine-Gray subdistribution hazard model has limitations. It retains individuals with competing events in the risk set, which is conceptually counterintuitive because these individuals are no longer at risk of the event of interest. While useful for prediction studies, the Fine-Gray model is not appropriate for causal or etiologic analyses. The model often requires large sample sizes to accurately estimate the subdistribution hazard of the event of interest and interpreting the subdistribution hazard ratios can be complex, particularly in comparison to cause-specific hazard ratios. The model also assumes proportional hazards for the subdistribution hazard, which might not always hold true in study populations. The cumulative incidence function estimated by the Fine-Gray model can sometimes overestimate the probability of the event of interest, particularly in the presence of strong competing risks.

In our example, cause-specific hazard ratios of the event of interest (MACE) and the competing event (non-cardiovascular death), along with corresponding 95% CIs, were estimated with Cox proportional hazards regression models. The cause-specific hazard ratios of MACE and non-cardiovascular death were 2.09 (95% CI: 2.06–2.13) and 3.35 (95% CI: 3.30–3.40), respectively. The subdistribution hazard ratio was estimated with the Fine-Gray competing risks regression model: the subdistribution hazard ratio of MACE in the VTE versus comparison cohorts was 1.48 (95% CI: 1.46–1.49).

Composite Endpoints and Competing Risks

To handle competing risks, a composite endpoint, such as MACE that includes acute myocardial infarction, stroke, and cardiovascular death,[24] is frequently used.[26] The composite endpoint occurs if any one of the defined components of the composite occurs.[26] The composite endpoint is therefore a single measure of effect that combines multiple individual endpoints into a single endpoint (Figure 5D). This approach is used to increase the statistical efficiency of a study by increasing the number of events and facilitating comparisons to be made between groups where the individual events are too rare to permit this. However, the individual components of a composite endpoint might not at all have the same clinical importance, plus the risk profile may be different for different components. Hence the interpretation of composite endpoints is not straightforward and even more so if one component of the composite endpoint is affected by competing risks.[26]

In our example, a composite endpoint of the outcome of interest (MACE) and the competing event (non-cardiovascular mortality) was specified, and the risk of this composite endpoint was estimated with the Kaplan-Meier method. The risk of the composite endpoint was 57.4% (95% CI: 57.2–57.7%) in the VTE cohort and 40.1% (95% CI: 39.9–40.3%) in the comparison cohort (Figure 6C).

Table 2 summarizes and compares the methodological approaches used to address competing risks in survival analysis. Figure 7 illustrates a decision tree that supports the choice of methods in the presence of competing risks.

Other Considerations

Common statistical software programs such as R, SAS, or Stata provide user-friendly macros and packages that make implementation of the presented methods used in competing risk analysis straightforward and accessible for researchers

Table 2 Comparison of Methodological Approaches to Competing Risks in Survival Analysis

	Aalen-Johansen Cumulative Incidence Estimation	**The Cause-specific Hazard Model**	**The Fine-Gray Subdistribution Hazard Model**	**Composite Endpoints**
Purpose	To directly estimate the cumulative incidence function in the presence of a competing event that precludes the occurrence of the event of interest	For etiologic research studies providing insight into mechanisms underlying each event type that may inform targeted public health interventions	To estimate the cumulative incidence of an event or predict a prognosis	In competing risk settings, composite endpoints help avoid bias that arises when competing events are censored.

(*Continued*)

Table 2 (Continued).

	Aalen-Johansen Cumulative Incidence Estimation	**The Cause-specific Hazard Model**	**The Fine-Gray Subdistribution Hazard Model**	**Composite Endpoints**
Approach	Non-parametric method that makes no assumptions about the underlying distribution of the time-to-event	Semi-parametric method that makes no assumptions about the distribution of the time-to -events Assumes competing risks are independent of each other and that hazards are proportional	Semi-parametric method that fits subdistribution hazard model Makes no assumptions about the distribution of the times-to-events Assumes proportional hazards for the subdistribution hazard	Including competing events as a component of a composite endpoint ensures they are properly accounted for in the analysis
Estimates	Directly estimates the cumulative incidence function, which represents the probability of specific events occurring at a given time, after allowing for the presence of a competing event	Provides cause-specific hazards	Estimates cumulative incidence function, the subdistribution hazards and the likelihood of an event after adjustment for covariates	The risk of the composite endpoint (outcome of interest and competing event combined) can be estimated with naïve Kaplan-Meier method.
Advantages	Avoids overestimation of the event probability that can occur with methods such as the naïve Kaplan-Meier estimator Provides a more valid estimate of cumulative incidence than other approaches such as Nelson-Aalen approach in the presence of competing risk	By modeling the hazard for each specific cause, possible to understand how each covariate affects the risk of each event separately Enables differentiation among various types of events (eg, different causes of death)	It is useful when the goal is to predict the absolute risk of an outcome event over time, rather than to study the cause or mechanism of the outcome event	Simple approach, easy to implement Reduces the number of pairwise comparisons, making results easier to interpret
Limitations	Unlike models like Fine-Gray, the Aalen-Johansen estimator does not allow covariates to be included The estimator assumes that censoring is independent of the event process. Violations of this assumption (eg, informative censoring) can lead to biased estimates	Gives cause-specific hazards but does not provide the probability of an outcome Can be difficult to interpret and may not be suitable for all study populations	Model often requires large sample sizes to give precise estimates for the subdistribution hazard of the event of interest. Interpreting the subdistribution hazard ratios can be complex, particularly in comparison to cause-specific hazard ratios Can sometimes overestimate the probability of the event of interest, particularly in the presence of strong competing risks	The interpretation of composite endpoints is complicated because individual components of a composite endpoint might not at all have the same clinical importance and the risk profile may be different for different components

(*In SAS, the PHREG procedure can fit both cause-specific Cox models and Fine-Gray subdistribution hazard models. In R, commonly used packages include cmprsk and riskRegression)*. In situations in which traditional methods are difficult to apply, when sample sizes are small, or the proportional hazards assumption might not hold, alternative solutions are

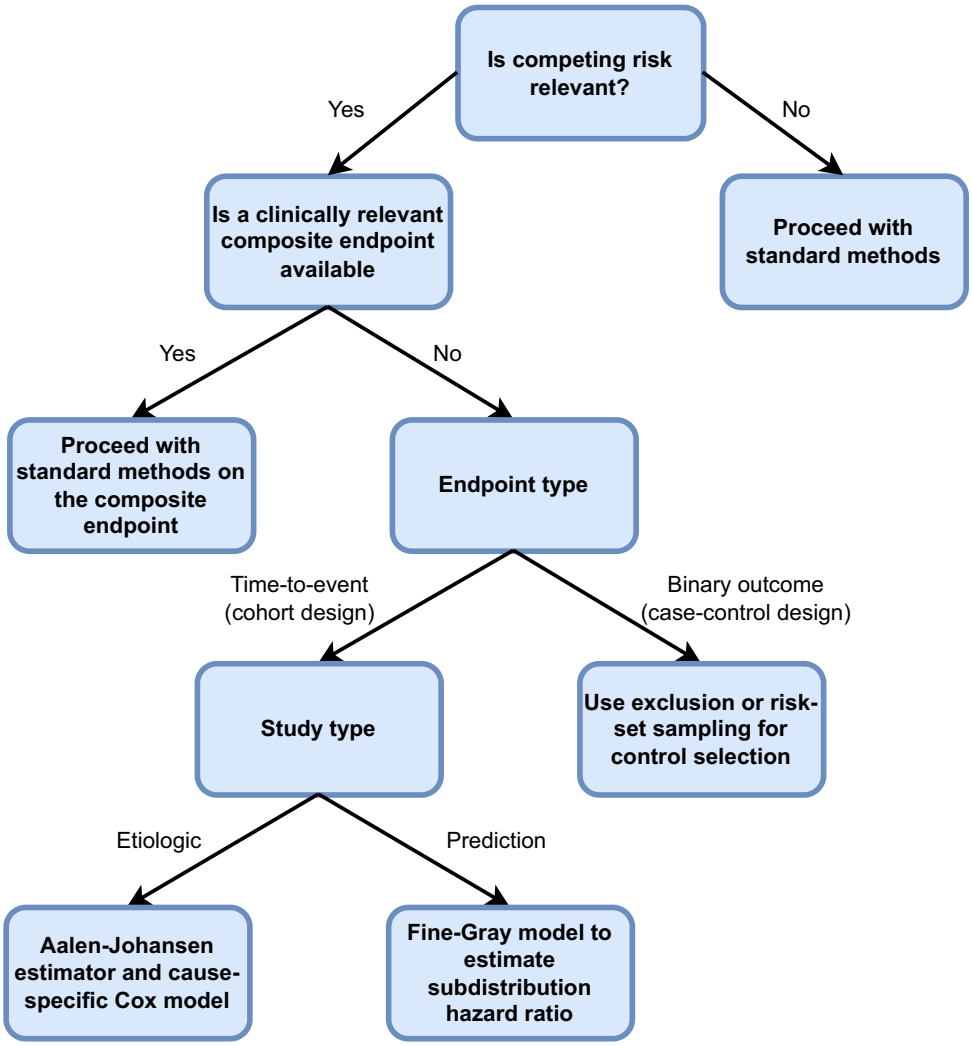

Figure 7 Decision tree illustrating choice of methods in the presence of competing risks.

available. The pseudo-value method involves computing pseudo-values from survival data and using these values in regression models.[27] This approach aids in providing more accurate estimates of survival probabilities. This method also allows for inclusion of time-dependent variables and interactions between covariates and therefore can be adapted to complex data structures. Statistical macros and functions are available to help the implementation of the pseudo-value method.[28]

Competing risks pose particular challenges in case-control studies because these designs estimate the odds ratio for a single outcome, whereas time-to-event analyses incorporate competing risks more naturally. Case-control studies examine antecedent exposures among cases compared with persons at risk of the illness. Incidence rates cannot typically be obtained for either exposed or non-exposed persons, but odds ratio can often be used to provide a valid estimate of the relative risk.

Competing risks, therefore, present particular challenges in case-control studies.

Several approaches can reduce the impact of competing risks in case-control studies.[29,30] For example, restriction can be applied by excluding individuals who experience a competing event before the index date, or controls can be matched on follow-up time to reduce bias. Risk set sampling, also called incidence density sampling, selects controls from individuals who remain at risk at the time each case occurs.[31] If a person experiences a competing event before the case index time, they are no longer in the risk set and cannot be selected as a control. Although risk set sampling does not fully model competing risks as the Fine–Gray model does, it reduces bias by excluding individuals who could never experience the outcome.

https://doi.org/10.2147/CLEP.S574928

In the classic cumulative case-control design, controls are sampled either from the entire source population at time of follow-up or, alternatively, from non-cases remaining at the end of follow-up after cases have been identified.[31] If the controls are samples from the non-cases at the end of follow-up, this design is called a cumulative case-control study. If the controls are sampled from the entire source population at start of follow-up, the design is called case-cohort design. Competing risks (eg, death from another cause before the outcome occurs) may influence the probability of developing the outcome over time.

The cumulative case-control design does not take competing risks into consideration. However, if the risk of a competing event is low, and the induction period short, the impact of competing risks is small.

We return to the use of prediction models in clinical care where a growing body of work has shown that even when rigorously validated, the models may not give accurate predictions. Problems can arise for a number of reasons including when current treatment at baseline is included as a predictor since this may be confounded with the indication. Further biases may result from known and unknown healthcare and secular changes that have taken place since the prediction model was developed. It is therefore recommended that the use of prediction models in clinical decision-making is monitored and that models are updated as needed with a stronger focus on causality.[32–34]

Methodological Advances

Other statistical approaches to modeling competing risks include analyzing cumulative incidence functions using non-parametric multiple imputation[35] and modeling the joint distribution of event time and event type.[36] Additional methods include multi-state models, which extend survival analysis by allowing individuals to transition between multiple health states, providing a flexible framework for analyzing complex event histories and competing risks,[37] and G-methods, which are specifically designed to handle time-varying confounders affected by prior treatment, thereby enabling valid causal inference in longitudinal studies.[38] The analysis of high-dimensional data brings particular computational challenges in the presence of competing risks and include penalized regression, boosting, random forest and deep learning. Further elaboration is beyond the scope of this paper but interested readers may find Djangang and colleagues arXiv paper that includes the mathematical details of these methods informative.[39]

Conclusion

The presence of competing risks poses an important challenge to valid estimation and interpretation of disease incidence and time-to-event data.[40] By using specialized statistical methods, such as cause-specific hazard models and subdistribution hazard models, more robust and more meaningful estimates of relative risks can be obtained. Understanding the effects of competing risks in the light of the specific study design and study purpose is crucial for developing effective public health and clinical interventions. Ignoring these risks can lead to biased results which may affect policy decisions and patient care.

Funding

Center for Population Medicine, Aarhus University Hospital and Aarhus University, Denmark.

Disclosure

Professor Henrik Sørensen reports funding for other studies in the form of institutional research grants to (and administered by) Aarhus University and Aarhus University confirms that none of these studies have any relation to the present study. He also reports fees for evaluation works from University of Oslo, the Norwegian Research Council, the Independent Research Fund Denmark, and the European Research Council. The authors report no other conflicts of interest in this work.

References

1. Hosmer DW, Lemeshow S, May S. Applied survival analysis: regression modeling of time-to-event data. In: *Wiley Series in Probability and Statistics*. Vol. xiii. 2nd. Wiley-Interscience; 2008:392.
2. Peacock JL, Peacock PJ. *Oxford Handbook of Medical Statistics*. Oxford Medical Publications. Oxford University Press; 2020.
3. Clayton D, Hills M. *Statistical Models in Epidemiology, Chapter 7: Competing Risks and Selection*. Oxford University Press; 2013.
4. Varadhan R, Weiss CO, Segal JB, Wu AW, Scharfstein D, Boyd C. Evaluating health outcomes in the presence of competing risks: a review of statistical methods and clinical applications. *Medical Care*. 2010;48(6 Suppl):S96–105. doi:10.1097/MLR.0b013e3181d99107

5. Partridge L, Deelen J, Slagboom PE. Facing up to the global challenges of ageing. *Nature*. 2018;561(7721):45–56. doi:10.1038/s41586-018-0457-8
6. Rothman KJ, Huybrechts KF, Murray EJ. *Epidemiology: An Introduction*. 3rd ed. Oxford University Press; 2024:286–288.
7. Clayton D, Hills M. *Statistical Models in Epidemiology. Statistical Models in Epidemiology, Chapter 27: Choice and Interpretation of Models*. Oxford University Press; 2013.
8. Koepsell TD, Weiss NS. *Epidemiologic Methods: Studying the Occurrence of Illness*. 2nd ed. Oxford University Press; 2014:55–56.
9. Katz MH.*Multivariable Analysis: A Practical Guide for Clinicians and Public Health Researchers*. Third Edition Vol. xv. Cambridge University Press;2011.233.
10. Sørensen ST, Kristensen FP, Troelsen FS, Schmidt M, Sørensen HT. Health registries as research tools: a review of methodological key issues. *Dan Med J*. 2023;70(4).
11. Parfrey P, Barrett B. *Clinical Epidemiology*. Vols. 155-62. 3rd. Humana Press; 2021
12. Groenwold RHH, le Cessie S, Dekkers OM. What is the research question? Estimands explained. *Eur J Endocrinol*. 2025;192(3):E5–e7. doi:10.1093/ejendo/lvaf048
13. Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: possibilities and pitfalls. *Int J Epidemiol*. 2012;41(3):861–870. doi:10.1093/ije/dyr213
14. Andersen PK, Abildstrom SZ, Rosthøj S. Competing risks as a multi-state model. *Stat Methods Med Res*. 2002;11(2):203–215. doi:10.1191/0962280202sm281ra
15. Sørensen HT, Horvath-Puho E, Pedersen L, Baron JA, Prandoni P. Venous thromboembolism and subsequent hospitalisation due to acute arterial cardiovascular events: a 20-year cohort study. *Lancet*. 2007;370(9601):1773–1779. doi:10.1016/s0140-6736(07)61745-0
16. Heide-Jørgensen U, Adelborg K, Kahlert J, Sørensen HT, Pedersen L. Sampling strategies for selecting general population comparison cohorts. *Clin Epidemiol*. 2018;10:1325–1337. doi:10.2147/clep.S164456
17. Schmidt M, Schmidt SAJ, Adelborg K, et al. The Danish health care system and epidemiological research: from health care contacts to database records. *Clin Epidemiol*. 2019;11:563–591. doi:10.2147/clep.S179083
18. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Stat Med*. 2007;26(11):2389–2430. doi:10.1002/sim.2712
19. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. *Circulation*. 2016;133(6):601–609. doi:10.1161/circulationaha.115.017719
20. Latouche A, Allignol A, Beyersmann J, Labopin M, Fine JP. A competing risks analysis should report results on all cause-specific hazards and cumulative incidence functions. *J Clin Epidemiol*. 2013;66(6):648–653. doi:10.1016/j.jclinepi.2012.09.017
21. Fine JP, Gray RJ. A proportional hazards model for the subdistribution of a competing risk. *J Ame Statist Assoc*. 1999;94(446):496–509. doi:10.1080/01621459.1999.10474144
22. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. *Am J Epidemiol*. 2009;170(2):244–256. doi:10.1093/aje/kwp107
23. Dayan V, Grant SW, Brophy JM, Barili F, Freemantle N. Composite end points and competing risks analysis. *Interdiscip Cardiovasc Thorac Surg*. 2024;39(1). doi:10.1093/icvts/ivae126
24. Satagopan JM, Ben-Porat L, Berwick M, Robson M, Kutler D, Auerbach AD. A note on competing risks in survival data analysis. *Br J Cancer*. 2004;91(7):1229–1235. doi:10.1038/sj.bjc.6602102
25. Austin PC, Fine JP. Practical recommendations for reporting fine-gray model analyses for competing risk data. *Stat Med*. 2017;36(27):4391–4400. doi:10.1002/sim.7501
26. Schulz KF, Grimes DA, Horton RC.*Essential Concepts in Clinical Research: Randomised Controlled Trials and Observational Epidemiology. P 200-202, and 208-209*. Second edition Vol. x. Elsevier;2019.256.
27. Andersen PK, Ravn H. Models for multi-state survival data: rates, risks, and pseudo-values. In: *Chapman & Hall/CRC Texts in Statistical Science Series*. CRC Press; 2024:278s.
28. Klein JP, Gerster M, Andersen PK, Tarima S, Perme MP. SAS and R functions to compute pseudo-values for censored data regression. *Comput Methods Programs Biomed*. 2008;89(3):289–300. doi:10.1016/j.cmpb.2007.11.017
29. Wolkewitz M, Cooper BS, Palomar-Martinez M, Olaechea-Astigarraga P, Alvarez-Lerma F, Schumacher M. Nested case-control studies in cohorts with competing events. *Epidemiology*. 2014;25(1):122–125. doi:10.1097/ede.0000000000000029
30. Hazard D, Schumacher M, Palomar-Martinez M, Alvarez-Lerma F, Olaechea-Astigarraga P, Wolkewitz M. Improving nested case-control studies to conduct a full competing-risks analysis for nosocomial infections. *Infect Control Hosp Epidemiol*. 2018;39(10):1196–1201. doi:10.1017/ice.2018.186
31. Rothman KJ, Huybrechts KF, Murray EJ. *Epidemiology: An Introduction*. 3rd ed. Oxford University Press; 2024:94–115.
32. Lenert MC, Matheny ME, Walsh CG. Prognostic models will be victims of their own success, unless…. *J Am Med Inform Assoc*. 2019;26(12):1645–1650. doi:10.1093/jamia/ocz145
33. van Amsterdam WAC, de Jong PA, Verhoeff JJC, Leiner T, Ranganath R. From algorithms to action: improving patient care requires causality. *BMC Med Inform Decis Mak*. 2024;24(1):111. doi:10.1186/s12911-024-02513-3
34. van Geloven N, Keogh RH, van Amsterdam W, et al. The risks of risk assessment: causal blind spots when using prediction models for treatment decisions. *Ann Intern Med*. 2025;178(9):1326–1333. doi:10.7326/annals-24-00279
35. Ruan PK, Gray RJ. Analyses of cumulative incidence functions via non-parametric multiple imputation. *Stat Med*. 2008;27(27):5709–5724. doi:10.1002/sim.3402
36. Nicolaie MA, van Houwelingen HC, Putter H. Vertical modeling: a pattern mixture approach for competing risks modeling. *Stat Med*. 2010;29(11):1190–1205. doi:10.1002/sim.3844
37. Moreno-Betancur M, Sadaoui H, Piffaretti C, Rey G. Survival analysis with multiple causes of death: extending the competing risks model. *Epidemiology*. 2017;28(1):12–19. doi:10.1097/ede.0000000000000531
38. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *BMJ*. 2017;359:j4587. doi:10.1136/bmj.j4587
39. Djangang PM, Han SS, Sanyal N. Comparative review of modern competing risk methods in high-dimensional settings. *arXiv:250312824v1 [statME]*. 2025.
40. Coemans M, Verbeke G, Döhler B, Süsal C, Naesens M. Bias by censoring for competing events in survival analysis. *BMJ*. 2022;378:e071349. doi:10.1136/bmj-2022-071349