Biomedical Statistics and Informatics
2026; 11(2): 60-64
http://www.sciencepublishinggroup.com/j/bsi
doi: 10.11648/j.bsi.20261102.12

Modelling Time-to-Failure in Presence of Competing Risks, the Sub-distribution Hazard Approach

Kipkirui Jephtah*, Tonui Benard Cheruiyot

Department of Mathematics, Actuarial Science & Statistics, University of Kabianga, Kericho, Kenya

Email address:

91rono@gmail.com (Kipkirui Jephtah), btonui@kabianga.ac.ke (Tonui Benard Cheruiyot)

To cite this article:

Kipkirui Jephtah, Tonui Benard Cheruiyot. (2026). Modelling Time-to-Failure in Presence of Competing Risks, the Sub-distribution Hazard Approach. *Biomedical Statistics and Informatics, 11*(2), 60-64. https://doi.org/10.11648/j.bsi.20261102.12

Received: 5 August 2025; **Accepted:** 27 November 2025 **Published:** 4 September 2026

Abstract: *Background:* Analysis of time-to-event data, generally called survival analysis, arise in many fields of study. Conventional methods of analyzing this type of data are historically well established and continue to be applied. The methods rely on the assumptions that subjects on follow-up can only fail from a well-defined single type of event, and that, conditioned on subject covariates, censoring time and the impending time-of-failure are independent. In real-world settings these assumptions are rarely fulfilled: subjects are typically exposed to multiple events that act concurrently to impede or modify the probability of failure from the event of interest. Events "compete" with each other, so the eventual failure of a subject can only be attributed to the first-occurring event – a competing risks scenario in which both the time to the first occurring event and the type of event that occurs at that time are of interest. *Objective:* To provide you the reader with a friendly and sufficient introduction to the theory underpinning the sub-distribution hazard model, based on the familiar or rather "traditional" description of survival data. This model describes the absolute risk of a subject experiencing an outcome while adjusting for right censoring, subject covariates, and naturally existing competing outcomes. *Conclusions:* In the competing risks framework, data may be modelled either through the cause-specific hazard model or the sub-distribution hazard model. Although both strategies are founded on the proportional hazards assumption, they differ in how risk sets are constituted and thus in what they measure, in their context of application, and in how the resulting hazard ratios are interpreted. For research questions where the sub-distribution hazard model is appropriate, the theoretical framework presented in this paper applies.

Keywords: Competing Risks, Survival Analysis, Censoring, Cumulative Incidence Function, Sub-distribution Hazard, Cause-specific Hazard

1. Introduction

The response variable of interest in survival analysis is the *time-to-occurrence* of a specific event, i.e. the time from a well-defined origin until a specified endpoint is reached. This time-to-event is often denoted by T and is interchangeably referred to as *survival time*, *event time*, or *"failure" time*. The distribution of T is conventionally characterized by its *survival function* $S(t) = Pr(T > t)$, the probability that an individual survives, or remains event-free, beyond a given time t.

Interest usually lies not in the distribution of survival times *per se*, but in how populations or sub-groups differ with respect to time of event occurrence given some exposure, intervention, or risk factor. A common difficulty is that not all subjects experience the event of interest by the close of a study: some are lost during follow-up because of some other, competing, event. This yields incomplete information about survival times, a phenomenon called *censoring*, and in particular *right censoring*.

According to [1], if censoring is not handled appropriately it has the potential to bias results and to reduce the statistical power of analyses. A clear understanding of *why* censoring

occurred is therefore essential for correct analysis of censored data.

1.1. Familiar or "Traditional" Description of Survival Data

Due to right censoring, the familiar or "traditional" description of survival data means that, for each i^{th} study subject $i = 1, 2, 3, ..., n$, we observe *identically, independently distributed* (i.i.d) random variables

$$\{T_i = min(T, C), \delta_i, E_i, \text{ and } X_i\}$$

where:

1) T_i is the potential true time of failure observed during the study;
2) $\delta_i = I(T \leq C)$ is a binary indicator of whether event time T or censoring time C was observed. If the event occurred ($T \leq C$), $\delta_i = 1$; if censoring occurred ($T > C$), $\delta_i = 0$. *Right-censoring* occurs when $T > C$, so $\delta_i = 0$ and consequently $(\delta_i, E_i) = 0$;
3) E_i denotes the event cause of failure (when $\delta_i = 1$);
4) X_i is a vector of, either continuous or categorical, covariates upon which survival time may depend.

In the survival literature, likelihood-based approaches to analyzing censored data (Kaplan–Meier curves, log-rank tests, and Cox regression) are well established and continue to be applied. Their success is hinged on two main assumptions: (1) that a subject can only fail from a single, well-defined event; and (2) that, given subject covariates, censoring occurs completely at random and independently of the event in question – i.e. censoring is *non-informative*, so censored and uncensored individuals have the same survival prospects. Non-informative censoring occurs when a subject is censored because the study period ends, the subject is lost to follow-up, or the subject fails due to something unrelated to the event of interest.

In real-world settings these two assumptions are rarely met, and careless application of the methods can lead to flawed results – for example, surrogate endpoints in AIDS treatment studies, where patients are withdrawn once CD4 counts fall below a threshold, produce clearly informative censoring.

1.2. Competing Risks/Events

According to [2], another difficulty commonly encountered in follow-up studies is the occurrence of a competing event/risk prior to the one of interest. A competing risk/event is defined as an event different from the one of interest that occurs first and impedes the event of interest from being observed, or fundamentally alters its probability of occurrence [3]. In this paper we presume competing risks/events are indexed $k = 1, 2, 3, ..., m$.

In the competing risks framework, data can be viewed as:

1) *Records arising from risk-specific latent failure times.* This approach was first developed by [4]. According to [5], latent failure times do not really exist, but are hypothetical times $\{T_{i1}, T_{i2}, ..., T_{im}\}$ that could be observed if a subject were to fail from each of the m causes operating in the population, so that $T_i = min\{T_{i1}, T_{i2}, ..., T_{im}\}$. The key additional assumption required is that causes of failure act independently on the population [6].
2) *Records of i.i.d. random variables* $\{T_i = min(T, C), \delta_i = I(T \leq C), E_i, X_i\}$, $i = 1, ..., n$. This second view permits modelling via estimable and identifiable functions, namely:
   a) *Cause-specific hazard modelling.* Every event/outcome has its own hazard function, so a usual Cox proportional hazards model can be fitted with respect to a specific outcome, hence *cause-specific hazard model*. According to [7], this can be used to assess covariate effects on a specific outcome, after treating individuals who fail from other causes as censored observations.
   b) *The sub-distribution hazard model.* An alternative that, besides adjusting for covariates, also accounts for competing outcomes directly by developing the sub-distribution hazard function, which quantifies the risk of failure from the k^{th} event given the subject has survived to time t *or* experienced a competing event $E \neq k$ at or before time t.[8]

The two hazard functions differ in how their *risk sets* are constituted. For the cause-specific hazard, subjects who experience the event of interest or a competing event at or before t are censored and removed from the risk set. For the sub-distribution hazard, the risk set consists of subjects who survived all events to time t *and* those who had a competing event at or before t – adjusted using *inverse probability of censoring weights* (Section 5.1).

2. The Model

Model (1) below, called the *sub-distribution hazard model*, was first developed by [8].

$$\begin{aligned} h_{ik}(t/X_i) &= h_{0k}(t) \exp\left(\beta_{k1}X_{i1} + \beta_{k2}X_{i2} + ... + \beta_{kp}X_{ip}\right) \\ &= h_{0k}(t) \exp\left(\boldsymbol{\beta}_k^T X_i\right) \end{aligned} \tag{1}$$

where $h_{ik}(t/X_i)$ is the predicted hazard at time t for the i^{th} individual to experience the k^{th} event, $h_{0k}(t)$ is the unspecified baseline sub-hazard function, and $\boldsymbol{\beta}_k^T X_i$ is the linear sum of covariate effects, taken to be *time-independent*. The model is a direct analogue of the [9] regression model, but with sub-distribution hazard functions in place of the traditional hazard functions.

The model stems from the *cumulative incidence function*

(CIF):

$$CIF_{ik}(t/X_i) = Pr(T \le t, E = k|X_i) = \int_0^t f_{ik}(u/X_i)du \quad \text{for } t > 0 \qquad (2)$$

which represents the absolute risk that the i^{th} individual experiences the k^{th} event as the first event, at time t, in presence of other causes. The CIF is also called the *sub-distribution function*: as $t \to \infty$, expression (2) converges to $Pr(E = k)$, not to 1 as in a proper probability distribution.

According to [8], the *sub-distribution hazard function* (SHR defined below) that uniquely describes the distribution in (2) is

$$h_{ik}(t/X_i) = \lim_{\triangle t \to 0} \frac{1}{\triangle t} Pr\Big(t \le T < t + \triangle t,\, E = k \mid \mathcal{R}_i(t)\Big) = \frac{f_{ik}(t)}{1 - CIF_{ik}(t)} = -\frac{d}{dt}\ln\left[1 - CIF_{ik}(t)\right] \qquad (3)$$

where $\mathcal{R}_i(t) = \{T > t\} \cup \{T \le t \text{ and } E \ne k\}$ is the conditioning event (survived to t, or previously had a competing event), $f_{ik}(t) = \frac{d}{dt}CIF_{ik}(t)$ is the sub-density function for the k^{th} event, and $1 - CIF_{ik}(t)$ represents the probability of surviving past t *or* having already failed from a competing cause $j \ne k$.

3. Research Gap

Leveraging the direct relationship between CIF (2) and the sub-hazard function (3),

$$CIF_{ik}(t/X_i) = 1 - \exp\left\{-\int_0^t h_{ik}(u/X_i)du\right\} \qquad (4)$$

A proportionality assumption was imposed on the sub-hazard functions, and the sub-distribution hazard model (1) was proposed under the counting process formulation of time-to-event data (see [10]) [8].

In this paper, we digress from the counting process formulation and instead explore the case when survival data is described in the familiar or "traditional" manner. To the best of our knowledge, no existing work has explored the Fine–Gray model under the traditional survival-data formulation, and this paper aims to fill that gap.

4. Model Assumptions

1) As with the [9] proportional hazards model, the ratio of sub-hazard functions comparing two individuals with different covariates is *proportional* and constant over time. If the sub-distribution hazard ratio is denoted SHR, then

$$SHR = \frac{h_{ik}(t/\,X_i)}{h_{jk}(t/\,X_j)} = \phi, \text{constant over time}$$

This means sub-hazard functions for different individuals, plotted on the complementary $log(-log)$ scale, should not cross.

2) For continuous covariates, the link function between the log of the sub-hazard ratio and the covariate vector is *linear*:

$$log\left[SHR(X_i)\right] = log\left\{\frac{h_{ik}(t/\,X_i)}{h_{ok}(t)}\right\} = (\beta_{k1}X_{i1} + \beta_{k2}X_{i2} + ... + \beta_{kp}X_{ip})$$

meaning a unit change in a continuous covariate has the same effect on the log-SHR regardless of covariate level.

5. Estimation of Model Parameters

Estimates of $\boldsymbol{\beta}_k^T$ for the sub-distribution hazard model (1) are obtained by *maximizing a partial likelihood function*, which has the same form as that of Cox's proportional hazards model.

According to [9], if one event occurs at time t, the conditional probability that it is from the i^{th} individual in the risk set $R(t)$ can be written as a regression model with $(p \times 1)$ vector $\boldsymbol{\beta}$. On cancelling the baseline hazard functions, the conditional probability that the failure observed at t is from individual i is

$$\frac{\exp(\boldsymbol{\beta}^T X_i)}{\sum\limits_{j\,\in\,R(t)} \exp(\boldsymbol{\beta}^T X_j)} \qquad (5)$$

Assuming individuals experience events independently (no ties), the joint likelihood over all n distinct event times, in terms of all survival times (censored or not), is

$$L(\boldsymbol{\beta}) = \prod_{i=1}^{n} \left\{ \frac{\exp\left(\boldsymbol{\beta}^T X_i\right)}{\sum\limits_{j \in R(t)} \exp\left(\boldsymbol{\beta}^T X_j\right)} \right\}^{\delta i} \qquad (6)$$

called the *partial likelihood function*, since it contains no information on the baseline hazard $h_o(t)$. Only subjects with $\delta_i = 1$ contribute; right-censored subjects ($\delta_i = 0$) do not.

Equation (6) was modified to accommodate m mutually exclusive events by: (1) redefining the risk set $R(t)$ to also include, as a placeholder, subjects who previously had a competing event at or before t, i.e. [8]

$$R(t) = \Big\{i; T > t \text{ or } \big(T \le t \text{ and } E \ne k\big)\Big\}$$

and (2) assigning such retained subjects time-dependent weights that gradually reduce according to the conditional probability of remaining under follow-up had the competing event not occurred – a procedure called *inverse probability of censoring weighting*.

5.1. Inverse Probability of Censoring Weights

These weights are

$$w_{ij}(t) = \frac{r_i(t)\hat{G}(t)}{\hat{G}\Big(\min(t, t_j)\Big)} \tag{7}$$
$$= \begin{cases} 1 & \text{if } T > t, \\ 0 & \text{if failed from event } k \text{ or censored at } t, \\ \frac{\hat{G}(t)}{\hat{G}(t_j)} \leq 1 & \text{if } T \leq t, E \neq k. \end{cases}$$

where $r_i(t) = I[C \geq min(T, t)]$ denotes vital status at time t. Censoring is assumed, per [11], to occur completely at random; the censoring distribution $\hat{G}(t) = Pr(C \geq t)$ is obtained via Kaplan–Meier by inverting event and censoring indicators. As noted by [12], the estimated probability of censoring is used to inversely weight the "complete data" score function. For a worked example, see chapter 6 of [13]. We concur with [14] that manual assignment of these weights is time-consuming; the *cmprsk* package's `crr()` function in *R* automates this. For a detailed, user-friendly implementation, see [15].

5.2. Likelihood Function

Implementing these modifications in (6) gives the partial likelihood function for all n individuals, with subjects in the risk set weighted:

$$L(\boldsymbol{\beta}_k) = \prod_{k=1}^{m} \left\{ \prod_{i=1}^{n} \left[\frac{\exp\left(\boldsymbol{\beta}_k^T X_i\right)}{\sum_{j \in R(t)} w_j(t)\ \exp\left(\boldsymbol{\beta}_k^T X_j\right)} \right]^{\delta_{ik}} \right\} \tag{8}$$

which, with respect to the k^{th} event type, factorizes into a product of m terms:

$$L(\boldsymbol{\beta}_k) = \prod_{i=1}^{n} \left[\frac{\exp\left(\boldsymbol{\beta}_k^T X_i\right)}{\sum_{j \in R(t)} w_j(t)\ \exp\left(\boldsymbol{\beta}_k^T X_j\right)} \right]^{\delta_{ik}} \tag{9}$$

Taking natural logarithms and simplifying gives the *log-partial likelihood function*:

$$\ln L(\boldsymbol{\beta}_k) = \sum_{i=1}^{n} \delta_{ik} \left\{ \boldsymbol{\beta}_k^T X_i - \ln \left[\sum_{j \in R(t)} w_j(t)\ \exp\left(\boldsymbol{\beta}_k^T X_j\right) \right] \right\} \tag{10}$$

Equation (10) is maximized, often via *Newton–Raphson iteration*, to yield $\hat{\beta}_{ks}$, the *maximum likelihood estimates* (MLE) of the true model parameters. If the series of iterations is $\beta^{\hat{j}+1}$ ($\hat{j} = 1, 2, 3, ...$), the iteration converges in probability to the true parameter values when $|\hat{\beta}^{\hat{j}+1} - \beta^{\hat{j}}| \approx 0$.

To permit further inference – Wald tests, likelihood ratio tests, and confidence intervals (CI) – It was shown that, under mild regularity conditions, the MLEs are *asymptotically multivariate normally distributed* [8]:

$$\sqrt{n}(\hat{\beta}_k - \beta_k) \approx N(0, \text{diagonal elements of } \mathbf{I}_{\beta_k}^{-1}), \quad \text{as } n \to \infty$$

where $\mathbf{I}_{\beta_k}$, the *Fisher Information matrix*, is a non-singular $(p \times p)$ matrix of negative expected second-order partial derivatives of $\ln L(\boldsymbol{\beta}_k)$. Its inverse is the asymptotic variance–covariance matrix of $\hat{\beta}_{ks}$; the square root of the j^{th} diagonal element is the standard error (s.e), giving, e.g., a 95% CI of $\hat{\beta}_k \pm 1.96\ s.e(\hat{\beta}_k)$.

Lastly, A proper way in which these estimated regression coefficients should be interpreted is described in [16].

6. Model Limitations

According to [17], one major limitation of the sub-distribution hazard model is that it requires the exact cause of failure to be known for every subject on follow-up.

7. Conclusion

Whereas censored survival times are inevitable in typical follow-up studies, competing events/risks occur naturally, and if ignored (by treating them as censored) there is a likelihood of inflating the risk of event occurrence – which, in clinical settings, is the cornerstone of decision-making. We conclude that when the sub-distribution hazard model is the appropriate choice for the research question at hand, the theory provided in this paper applies.

ORCID

0009-0005-2597-1342 (Kipkirui Jephtah)
0009-0009-1630-3827 (Tonui Benard Cheruiyot)

Abbreviations

CIF	Cumulative Incidence Function
SHR	Sub-distribution Hazard Ratio
MLE	Maximum Likelihood Estimate(s)
CI	Confidence Interval
i.i.d	Identically, Independently Distributed
s.e	Standard Error
ORCID	Open Researcher and Contributor ID

Author Contributions

Kipkirui Jephtah: Funding Acquisition, Supervision, Writing – review & editing

Tonui Benard Cheruiyot: Funding Acquisition, Supervision, Writing – review & editing

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Turkson, A. J., Ayiah-Mensah, F. and Nimoh, V. (2021). Handling Censoring and Censored Data in Survival Analysis: A Standalone Systematic Literature Review. *International Journal of Mathematics and Mathematical Sciences*, 2021, 1–16. https://doi.org/10.1155/2021/9307475

[2] Moeschberger, M. L. and Klein, J. P. (1995). Statistical Methods for Dependent Competing Risks. *Lifetime Data Analysis, 1, 195–204* https://doi.org/10.1007/BF00985770

[3] Gooley, T. A., Leisenring, W., Crowley, J. and Storer, B. E. (1999). Estimation of Failure Probabilities in the Presence of Competing Risks: New Representations of Old Estimators. *Statistics in Medicine*, 18(6), 695–706. https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<695::AID-SIM60>3.0.CO;2-O

[4] Cox, D. R. (1959). The Analysis of Exponentially Distributed Life-Times with Two Types of Failure. *Journal of the Royal Statistical Society: Series B*, 21(2), 411–421. https://doi.org/10.1111/j.2517-6161.1959.tb00349.x

[5] Elandt-Johnson, R. C. and Johnson, N. L. (1980). *Survival Models and Data Analysis*, Vol. 110. John Wiley & Sons, New York.

[6] Gail, M. (1975). A Review and Critique of Some Models Used in Competing Risk Analysis. *Biometrics*, 31(1), 209–222. https://doi.org/10.2307/2529721

[7] Prentice, R. L., Kalbfleisch, J. D., Peterson Jr., A. V., Flournoy, N., Farewell, V. T. and Breslow, N. E. (1978). The Analysis of Failure Times in the Presence of Competing Risks. *Biometrics*, 34(4), 541–554. https://doi.org/10.2307/2530374

[8] Fine, J. P. and Gray, R. J. (1999). A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*, 94(446), 496–509. https://doi.org/10.1080/01621459.1999.10474144

[9] Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society: Series B*, 34(2), 187–202. https://doi.org/10.1111/j.2517-6161.1972.tb00899.x

[10] Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (2012). *Statistical Models Based on Counting Processes*. Springer Science & Business Media, New York. https://doi.org/10.1007/978-1-4612-4348-9

[11] Collett, D. (2015). *Modelling Survival Data in Medical Research* (3rd ed.). CRC Press. https://doi.org/10.1201/9781003282525

[12] Bajorunaite, R. and Klein, J. P. (2008). Comparison of Failure Probabilities in the Presence of Competing Risks. *Journal of Statistical Computation and Simulation*, 78(10), 951–966.

[13] Pintilie, M. (2006). *Competing Risks: A Practical Perspective*. John Wiley & Sons, Chichester. https://doi.org/10.1002/9780470870709

[14] Donoghoe, M. W. and Gebski, V. (2017). The Importance of Censoring in Competing Risks Analysis of the Subdistribution Hazard. *BMC Medical Research Methodology*, 17, 52. https://doi.org/10.1186/s12874-017-0327-3

[15] Willems, S. J. and Fiocco, M. (2014). Inverse Probability Censoring Weights for Routine Outcome Monitoring Data. Technical Report, Universiteit Leiden, Leiden, The Netherlands. https://hdl.handle.net/1887/3597243

[16] Austin, P. C. and Fine, J. P. (2017). Practical Recommendations for Reporting Fine-Gray Model Analyses for Competing Risk Data. *Statistics in Medicine*, 36(27), 4391–4400. https://doi.org/10.1002/sim.7501

[17] Mao, L. and Lin, D. Y. (2017). Efficient Estimation of Semiparametric Transformation Models for the Cumulative Incidence of Competing Risks. *Journal of the Royal Statistical Society: Series B*, 79(2), 573–587. https://doi.org/10.1111/rssb.12177