Survival Analysis and Frailty Models: Applications with R

Claudia Adriana Castro Kuriss
Instituto Tecnológico de Buenos Aires, Argentina

Abstract

Survival analysis is increasingly utilized in diverse fields such as health sciences, engineering, social sciences, and psychology. The objective is to study subjects and record the time until an event of interest occurs, such as an infection, a transplant, an engine failure, a psychiatric patient relapse, or obtaining a new employment. When the event does not occur, the associated times are considered censored. There is a growing need for more flexible models beyond the widely known Cox proportional hazards model. These include models that incorporate time-dependent covariates and multi-state models. Among these models are those for recurrent events, which account for repeated occurrences of the event over time, such as multiple re-hospitalizations, multiple gout attacks, or disease exacerbations before death. Another important category is competing risks models, where multiple events compete to be the primary outcome, such as different causes of death. Additionally, frailty models address the issue of certain individuals or groups being more vulnerable or frail than others. These models consider unobserved heterogeneity among individuals or clusters, such as families, countries, or hospitals, where survival times within a cluster are similar. These individuals are said to share frailty. Frailty would account for the observed heterogeneity between individuals and would be a multiplicative random effect in the model, either for single individuals or groups. In this work, we will explore techniques for fitting both univariate and multivariate frailty models using the R programming language. We will assume the frailty variable follows a parametric distribution and will examine the proposed methods for addressing these problems. The developed theory will be applied to a real medical example.

1 Introduction

Survival analysis provides a fundamental set of tools for studying time-to-event data and is widely applied across a broad range of scientific disciplines, including health sciences, engineering, social sciences, and psychology. Typical objectives include modeling the time until infection or death, disease relapse, system failure, hospital readmission, time to retirement, often in the presence of censoring and truncation. In many applied settings, however, survival data exhibit structural complexities that challenge standard modeling approaches.

A common feature of contemporary applied analyses is the presence of multiple sources of dependence and heterogeneity. Event times may be correlated due to repeated occurrences within individuals, clustering within families, hospitals, or regions, or the coexistence of non-terminal and terminal events. Moreover, individuals with similar observed covariates may experience markedly different outcomes due to latent factors that are not captured in the available data. Despite these complexities, applied analyses frequently proceed by selecting a convenient model to fit a given dataset, with limited scrutiny of the assumptions underlying the chosen formulation or of the consequences of model misspecification.

The Cox proportional hazards model (Cox, 1972) remains the most widely used framework in survival analysis. While its semi-parametric nature and interpretability make it attractive in many contexts, the Cox model may be inadequate when unobserved heterogeneity is present or when the event process departs from a single, independent failure time. In such cases, ignoring latent heterogeneity or dependence structures can lead to biased regression estimates, distorted marginal hazard functions, and misleading substantive conclusions.

Frailty models were introduced as a principled extension of proportional hazards models to account for unobserved heterogeneity by incorporating a latent random effect into the hazard function. This frailty term acts multiplicatively on the baseline hazard and represents unmeasured vulnerability or susceptibility at the individual or group level. By inducing dependence between event times, frailty models provide a natural framework for recurrent events, clustered survival data, and other settings where independence assumptions are untenable. Importantly, the distinction between conditional (frailty-specific) and marginal (population-level) quantities highlights how unobserved heterogeneity fundamentally alters survival and hazard behavior over time (Hougaard, 2012; Duchateau and Janssen, 2010).

In parallel, multi-state and competing risks models have been developed to analyze complex event histories involving transitions between multiple states, such as illness, recovery, relapse, and death (Putter et al., 2007). These models emphasize the dynamic structure of event processes but typically assume that, conditional on observed covariates, individuals are homogeneous. Frailty and multi-state modeling therefore address complementary aspects of survival data, and their integration has become increasingly relevant in modern applications (Hougaard, 1999).

A particularly challenging setting arises when recurrent non-terminal events are observed in the presence of a terminal event that precludes further observation, such as hospital readmissions followed by death. In such semi-competing risks or illness–death frameworks without recovery, the recurrent and terminal processes are typically dependent, and separate analyses may obscure their association. Joint frailty models provide a coherent methodological solution by linking the recurrent and terminal event intensities through a shared latent frailty term and, in some formulations, an explicit association parameter.

The primary objective of this paper is methodological. We provide an integrated overview of frailty models and their extensions, with particular emphasis on recurrent events, semi-competing risks, and joint frailty formulations. Beyond reviewing model structures, we discuss key inferential issues, including the interpretation of conditional versus marginal effects, the role of frailty distributional assumptions, and practical challenges related to identifiability and numerical stability.

To illustrate these methodological aspects, we consider data from a cohort of patients undergoing cardiac surgery. This dataset is used not as an end in itself, but as a vehicle to demonstrate how different frailty-based modeling choices affect inference in realistic settings, and to highlight practical issues that arise when applying existing software to complex event processes. Through this illustrative example, we aim to clarify both the strengths and limitations of joint frailty models in applied survival analysis.

2 Frailty Models: framework

Frailty models extend classical survival models by introducing a latent random effect that accounts for unobserved heterogeneity among individuals or clusters. The key idea is that individuals with similar observed covariates may still exhibit different risk profiles due to unmeasured or unobservable factors. These latent effects are summarized through a frailty term, which modifies the hazard function and induces dependence between event times.

Let T denote a non-negative survival time and let $\mathbf{X}$ be a vector of observed covariates. In a proportional hazards framework with frailty, the conditional hazard function for an individual given a frailty value $Z = z$ is defined as

$$\lambda(t \mid Z = z, \mathbf{X}) = z\, \lambda_0(t) \exp(\boldsymbol{\beta}^\top \mathbf{X}), \tag{1}$$

where $\lambda_0(t)$ is the baseline hazard function and $\boldsymbol{\beta}$ is the vector of regression coefficients. The frailty term Z is assumed to be a non-negative random variable with $\mathbb{E}(Z) = 1$ and variance $\mathrm{Var}(Z) = \theta$, to be estimated from the data. The normalization $\mathbb{E}(Z) = 1$ ensures identifiability of the baseline hazard.

Conditional on Z, individuals are assumed to be independent and to follow a standard proportional hazards model. The conditional survival function is therefore given by

$$S(t \mid Z = z, \mathbf{X}) = \exp\Big\{-z\, \Lambda_0(t) \exp(\boldsymbol{\beta}^\top \mathbf{X})\Big\}, \tag{2}$$

where $\Lambda_0(t) = \int_0^t \lambda_0(u)\, du$ denotes the cumulative baseline hazard.

In practice, interest often lies in marginal (population-level) quantities rather than conditional ones. The marginal survival function is obtained by integrating the conditional survival function with respect to the frailty distribution,

$$S(t \mid \mathbf{X}) = \mathbb{E}_Z\left[S(t \mid Z, \mathbf{X})\right] = \mathbb{E}_Z\left[\exp\Big\{-Z\, \Lambda_0(t) \exp(\boldsymbol{\beta}^\top \mathbf{X})\Big\}\right]. \tag{3}$$

The expression in (3) highlights the central role of the Laplace transform of the frailty distribution. Let $\mathcal{L}_Z(s) = \mathbb{E}[\exp(-sZ)]$ denote the Laplace transform of Z. Then the marginal survival function can be written compactly as

$$S(t \mid \mathbf{X}) = \mathcal{L}_Z\Big(\Lambda_0(t) \exp(\boldsymbol{\beta}^\top \mathbf{X})\Big). \tag{4}$$

As a consequence, frailty distributions with tractable Laplace transforms lead to closed-form expressions for marginal survival and hazard functions, which is one of the main reasons why certain distributions are preferred in applied work.

A direct implication of unobserved heterogeneity is the distinction between conditional and marginal hazards. While the conditional hazard in (1) follows a proportional hazards structure, the marginal hazard generally does not. At the population level, frailty induces a form of selection over time, as more vulnerable individuals tend to experience the event earlier, leaving a progressively more robust subset of survivors. This phenomenon explains why marginal survival curves typically decline faster than their conditional counterparts and why population hazards may exhibit non-proportional behavior even when the conditional model is proportional.

Frailty models can be formulated at different levels. In univariate frailty models, the frailty term is associated with a single survival time per individual. In multivariate or clustered settings, a shared frailty term is introduced to model dependence between multiple event times within the same cluster, such as repeated events within an individual or correlated outcomes within families or hospitals. More complex extensions include joint frailty models, which link recurrent events with a terminal event through a common frailty term and, in some cases, an additional association parameter.

Frailty models have been extensively studied and developed as a flexible framework for modeling unobserved heterogeneity in survival data (Wienke, 2010).

This general framework provides the basis for the different frailty distributions and model classes discussed in the following sections.

2.1 Interpretation of Frailty Effects

Frailty acts as a latent multiplicative effect on the hazard function and induces heterogeneity in individual risk profiles. From a population perspective, this heterogeneity leads to distinct marginal survival and hazard patterns compared to models without frailty. Figure 1 illustrates the impact of frailty on baseline survival and hazard functions for different levels of unobserved heterogeneity.

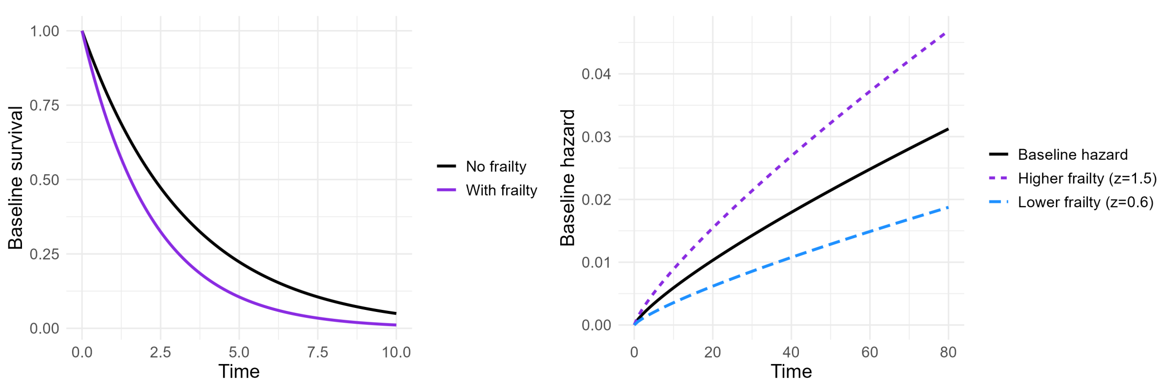

Figure 1: Effect of frailty on baseline survival (left) and baseline hazard functions (right). Unobserved heterogeneity shifts the baseline survival downward, reflecting earlier failures among more vulnerable individuals, and modifies the baseline hazard depending on the magnitude of the frailty term.

3 Frailty Distributions

The choice of the frailty distribution plays a central role in both the theoretical properties and the practical implementation of frailty models. Since marginal survival and hazard functions are obtained by integrating out the latent frailty term, tractability of this integration is a key consideration. In particular, distributions whose Laplace transform is available in closed form lead to explicit expressions for marginal quantities and facilitate likelihood-based inference.

Let Z denote a non-negative frailty variable with mean $\mathbb{E}(Z) = 1$ and variance $\text{Var}(Z) = \theta$. As shown in Section 2, the marginal survival function depends on the Laplace transform of Z, evaluated at the cumulative hazard. Consequently, much of the literature has focused on frailty distributions for which the Laplace transform has a simple analytic form.

The Gamma family is the predominant choice for modeling frailty. Its popularity stems from both mathematical convenience and interpretability. The Gamma distribution belongs to the class of infinitely divisible distributions, which implies that it can be represented as the sum of an arbitrary number of independent and identically distributed random variables. This property leads to a simple Laplace transform and allows the marginal likelihood to be expressed in closed form under proportional hazards assumptions. As a result, Gamma frailty models are computationally stable and have been extensively applied in medical and epidemiological studies.

Another commonly used choice is the inverse Gaussian frailty distribution, which also admits a closed-form Laplace transform and belongs to the infinitely divisible family. Inverse Gaussian frailty induces a different form of heterogeneity and may be more appropriate in settings where the tail behavior of the frailty distribution differs from that implied by the Gamma model.

The log-normal distribution has also been proposed as a frailty distribution, particularly in multivariate and correlated frailty settings. Unlike the Gamma and inverse Gaussian cases, the log-normal distribution is not infinitely divisible and does not admit a closed-form Laplace transform. As a consequence, likelihood-based inference typically relies on numerical integration or approximation methods, which may increase computational complexity and affect numerical stability. Despite these limitations, log-normal frailty offers greater flexibility in modeling dependence structures in some multivariate contexts.

More general families of frailty distributions have been developed to encompass a wider range of dependence patterns. These include positive stable distributions and the Power-Variance Function (PVF) family, which contains the Gamma distribution as a special case. Hougaard distributions provide a unifying framework that allows different frailty structures to be obtained through appropriate parameter choices, offering additional flexibility while preserving analytic tractability in many cases, (Hougaard, 1986).

More recently, alternative distributions such as reparametrized Birnbaum–Saunders frailty have been proposed, (Leão et al., 2018). Although this distribution does not belong to the infinitely divisible family, it possesses a tractable Laplace transform, making it suitable for likelihood-based inference. Such distributions may be particularly appealing in applications where the frailty mechanism is associated with fatigue or aging-type processes.

Recent work has compared estimation strategies for shared frailty models, highlighting differences in finite-sample performance and numerical stability (Wu et al., 2023).

It is important to emphasize that there is no universally optimal choice of frailty distribution. Different distributions imply different forms of unobserved heterogeneity and lead to distinct marginal hazard and survival behaviors. In practice, the choice of frailty distribution should be guided by a combination of theoretical considerations, computational feasibility, and empirical model fit. In the application presented in this work, several frailty distributions are examined and compared using likelihood-based criteria in order to assess their adequacy for the data at hand,(Hougaard, 2012; Duchateau and Janssen, 2010; Wienke, 2010).

4 Model Classes

Frailty models can be formulated in a variety of ways depending on the structure of the data and the scientific question of interest. The main distinction concerns whether the frailty term is associated with a single survival time, multiple correlated event times, or both recurrent and terminal events. In this section, we briefly review the most common classes of frailty models used in practice.

4.1 Univariate Frailty Models

The simplest setting corresponds to univariate frailty models, where a single event time is observed for each individual. In this case, the frailty term captures unobserved heterogeneity across individuals that cannot be explained by observed covariates. The model assumes conditional proportional hazards given the frailty, and marginal inference is obtained by integrating out the frailty distribution.

Univariate frailty models can be specified using parametric baseline hazards, such as the exponential or Weibull distributions, or within a semi-parametric framework that leaves the baseline hazard unspecified. These models are appropriate when individuals can be viewed as an independent random sample from the population and when unobserved individual-level risk variation is suspected.

4.2 Shared Frailty Models for Clustered Data

In many applications, survival times are correlated because individuals belong to common clusters, such as families, hospitals, or geographical regions. Shared frailty models extend the univariate framework by introducing a common frailty term for all individuals within the same cluster. Conditional on the shared frailty, event times are assumed to be independent.

This formulation induces dependence between survival times within clusters and provides a parsimonious way to model within-cluster correlation. Shared frailty models have been widely used for clustered survival data and recurrent event settings, where repeated events within the same individual are treated as belonging to the same cluster.

4.3 Frailty Models for Recurrent Events

Recurrent event data arise when the same type of event can occur repeatedly for the same individual over time, such as hospital readmissions or repeated disease relapses. Frailty models provide a natural framework for such data by introducing a subject-specific frailty term that is shared across all recurrent events for an individual.

This shared frailty induces dependence between successive events and accounts for unobserved individual-level risk factors. Conditional on the frailty, recurrent events are assumed to follow a proportional intensity model. Marginally, the frailty induces overdispersion and correlation among event times.

4.4 Joint Frailty Models and Semi-Competing Risks

In many medical applications, recurrent events are observed in the presence of a terminal event, such as death, which precludes further observation of recurrent events. In this context, the recurrent event process and the terminal event are typically dependent. Joint frailty models address this dependence by linking both processes through a shared frailty term and, in some formulations, an additional association parameter.

A common special case is the illness–death model without recovery, which can be viewed as a semi-competing risks framework. In this setting, the non-terminal event (e.g., hospital readmission) may occur before the terminal event (e.g.,death), but not vice versa. Joint frailty models provide a coherent way to analyze such data by accounting simultaneously for unobserved heterogeneity and dependence between the non-terminal and terminal events.

4.5 Other Extensions

More complex frailty structures have been proposed to address additional sources of dependence. Correlated frailty models allow frailty terms associated with different individuals within a cluster to be correlated, as in twin or sibling studies. Additive frailty models relax the multiplicative assumption by allowing frailty effects to enter the hazard additively. Nested frailty models accommodate hierarchical clustering structures with multiple levels of unobserved heterogeneity.

Although these extensions offer greater modeling flexibility, they also increase computational complexity and may pose challenges in terms of identifiability and numerical stability. Consequently, their use is typically motivated by specific scientific questions or data structures.

The choice among different frailty model classes should be guided by the nature of the event process, the presence of clustering or recurrence, and the role of terminal events. In the following sections, we focus on estimation methods and software implementations for the model classes most relevant to the application considered in this study.

5 Estimation Methods and Software Implementation

Inference in frailty models is typically based on likelihood-based approaches that integrate out the latent frailty term. The choice of estimation method depends on the assumed frailty distribution, the structure of the baseline hazard, and the complexity of the event process (e.g., recurrent events, terminal events, or joint models). In this section, we review the main estimation strategies commonly used in practice and discuss their implementation in available R packages. Several R packages implementing frailty models are explored in this study and all analyses are conducted using `R` (R Core Team, 2025).

5.1 Marginal Likelihood via Laplace Transform

When the frailty distribution admits a closed-form Laplace transform, the marginal likelihood can be obtained analytically by integrating out the latent frailty variable. This approach is most commonly used with Gamma frailty, where the frailty variable Z satisfies $\mathbb{E}(Z) = 1$ and $\mathrm{Var}(Z) = \theta$. For clustered or recurrent event data, the marginal likelihood contribution of a cluster can be expressed explicitly in terms of the Laplace transform of the frailty distribution.

Maximization of the marginal log-likelihood is typically carried out using Newton–Raphson or related optimization algorithms. Standard errors of the parameter estimates are obtained from the inverse of the observed Hessian matrix. This estimation strategy is implemented in `R`, in the `survival` package through the `coxph` function with a frailty term,(Therneau, 2023), as well as in `frailtypack` for more general model structures, (Rondeau et al., 2012). While this approach is computationally efficient and widely used, it is limited to frailty distributions with tractable Laplace transforms and does not naturally extend to all joint frailty settings.

5.2 Penalized Likelihood and Laplace Approximation

When a closed-form Laplace transform is not available, likelihood-based inference can be performed using penalized likelihood methods combined with numerical approximation. In this framework, the baseline hazard function is modeled flexibly using splines, and a penalty term is introduced to control smoothness. Integration over the frailty distribution is then approximated using a Laplace approximation.

This approach allows for greater flexibility in both the baseline hazard and the choice of frailty distribution, including inverse Gaussian and other non-Gamma frailties. Penalized likelihood methods are implemented in the `frailtypack` package and support a wide range of models, including recurrent events, terminal events, and joint frailty models. However, increased flexibility may come at the cost of higher computational complexity and potential numerical instability in finite samples.

5.3 Pseudo Full Likelihood Approaches

Pseudo full likelihood (PFL) methods provide a semi-parametric alternative that avoids explicit specification of the baseline hazard function. In this framework, the baseline hazard is treated as a nuisance parameter and is profiled out using a Breslow-type estimator, (Breslow, 1972),while inference on the regression coefficients and frailty variance is based on likelihood-type score equations.

Under suitable regularity conditions, PFL estimators are consistent and asymptotically normal. This approach is particularly attractive in shared frailty settings with recurrent events and has been implemented in the `frailtySurv` package, which supports several frailty distributions, including Gamma, log-normal, inverse Gaussian, and power-variance function families, (Monaco et al., 2018). The semi-parametric nature of this method provides robustness with respect to baseline hazard misspecification, although interpretation and computation may be more involved.

5.4 Expectation–Maximization Algorithms

Expectation–Maximization (EM) algorithms treat the frailty terms as latent variables and alternate between computing their conditional expectations given the data (E-step) and maximizing the expected complete-data log-likelihood with respect to the model parameters (M-step). This approach allows for flexible modeling of frailty distributions and can be applied in a variety of settings.

The EM algorithm is implemented in the `frailtyEM` package, which supports several frailty distributions and semi-parametric baseline hazards, (Balan and Putter, 2019). Although EM-based methods are appealing due to their generality, they may suffer from slow convergence or instability of variance estimates, particularly in datasets with limited information on the frailty parameters.

5.5 Mixed-Effects Model Analogy

An alternative perspective treats frailty as a random effect within a mixed-effects modeling framework. In this approach, frailty is modeled as a random intercept, and estimation is based on penalized partial likelihood. This formulation establishes a close connection between frailty models and generalized linear mixed models.

The mixed-effects analogy is implemented in the `coxme` package, which performs well for clustered survival data with large cluster sizes, (Therneau, 2024). However, this approach is limited in its ability to handle terminal events jointly with recurrent events and is therefore not suitable for semi-competing risks or joint frailty models.

5.6 Practical Implementation and Software Considerations

In applied analyses, implementation details play a crucial role in the performance and interpretability of frailty models. Different R packages require different data structures, such as long or start–stop formats, cluster indicators, or transition-specific datasets in multi-state and joint frailty settings. Consequently, each dataset needs to be tailored to the specific requirements of each estimation framework.

In addition, modifications to the default implementation may be necessary in order to ensure consistency with the underlying modeling assumptions. Observed issues such as near-singular Hessian matrices or boundary estimates of frailty variance should be interpreted as indications of weak identifiability or flat likelihood surfaces in finite samples, rather than as failures of the frailty modeling framework itself.

Overall, no single estimation method or software package provides a universal solution. Instead, reliable inference in frailty models requires a careful combination of appropriate model specification, data preparation, and diagnostic assessment of numerical stability.

6 Application to Cardiac Surgery Data

6.1 Illustrative Example: study design and data structure

To illustrate the methodological issues discussed in the previous sections, we consider data from a cohort of patients who underwent cardiac surgery in a hospital in Buenos Aires, Argentina. The purpose of this example is not to provide a comprehensive clinical analysis, but rather to demonstrate how frailty-based modeling choices affect inference in the presence of recurrent non-terminal events and a terminal event.

The study included 1,327 patients operated on during 2014 and followed for up to 1,436 days. The mean age at surgery was 66 years (standard deviation 11.27 years), and 74% of the patients were male. Surgical procedures comprised myocardial revascularization (48.9%), valve surgery (22%), and combined procedures.

Baseline covariates included demographic, clinical, and laboratory variables routinely collected in clinical practice, such as age, sex, body mass index, atrial fibrillation, diabetes, smoking status, respiratory history, dyslipidemia, arterial hypertension, urea level, hematocrit, corrected creatinine, use of extracorporeal circulation, and insulin treatment. These covariates are used to illustrate how joint frailty models accommodate a rich set of predictors while accounting for unobserved heterogeneity and dependence between event processes.

The events of interest were hospital readmissions following discharge and death. Readmissions may occur repeatedly over time, whereas death constitutes a terminal event that precludes further observation of readmissions. This structure naturally motivates a modeling framework that jointly accounts for recurrent events, terminal events, and latent patient-specific heterogeneity.

6.2 Illness–death framework and joint frailty specification

The event history described above was analyzed within an illness–death framework without recovery, corresponding to a semi-competing risks setting. Because all patients were discharged alive after surgery, this outcome is classified as state 0. The remaining states are defined as follows: state 1 (hospital readmission) and state 2 (death). Possible transitions include $0 \to 1$ (readmission), $0 \to 2$ (death without prior readmission), and $1 \to 2$ (death after readmission).

Let $\lambda_{01}(t)$, $\lambda_{02}(t)$, and $\lambda_{12}(t)$ denote the transition-specific hazard functions. For each transition, a proportional hazards structure with a shared subject-specific frailty term was assumed. The frailty term acts multiplicatively on all transition intensities, capturing unobserved patient-level susceptibility and inducing dependence between the recurrent readmission process and the terminal event.

This joint frailty formulation provides a coherent approach to semi-competing risks data, where the non-terminal event may occur prior to death but not vice versa. In contrast to separate analyses of readmission and mortality, the joint model explicitly represents their association through a common latent component, allowing both processes to be interpreted conditionally on frailty and marginally at the population level.

6.3 Estimation Strategy and Software Considerations

Continuous covariates were rescaled for interpretability and numerical stability. In particular, some variables were divided by 10 so that hazard ratios correspond to clinically meaningful changes, while selected covariates were standardized to facilitate comparison of effect sizes across predictors.

Several R packages implementing frailty models were explored in preliminary analyses in order to assess both inferential consistency and numerical stability in this complex setting. Estimation based on the `frailtyEM` package frequently exhibited convergence difficulties or near-singular Hessian matrices, likely reflecting limited information on variance and association parameters in finite samples. The `frailtySurv` package provided stable estimates of frailty variance for recurrent event models but did not support joint modeling of recurrent and terminal events.

Analyses conducted using the mentioned `survival` and `coxme` packages were restricted to the readmission process and therefore did not allow direct assessment of dependence between readmission and death, although estimated covariate effects were broadly consistent with those obtained under joint modeling.

Final analyses were therefore performed using the `frailtypack` package, which supports joint frailty models for recurrent and terminal events within a unified framework. Our dataset was adapted to meet the specific structural requirements of each implementation, and also, some adjustments were introduced to ensure that likelihood contributions were consistent with the semi-competing risks design and did not artificially accumulate events across model components, which may lead to unstable parameter estimates.

This comparative implementation highlights a key methodological point: while multiple software tools exist for frailty modeling, their scope, assumptions, and numerical behavior differ substantially. Careful alignment between the scientific question, the chosen model formulation, and the software implementation is therefore essential for reliable inference.

6.4 Frailty Distribution and Model Selection

Several frailty distributions were evaluated, including Gamma, inverse Gaussian, log-normal, and stable distributions. Model adequacy was assessed using penalized marginal log-likelihood, likelihood cross-validation criteria, and information criteria such as AIC and BIC. Across both single-covariate and multivariable models, the Gamma frailty distribution consistently provided the best fit and stable parameter estimates.

6.5 Results

6.5.1 Univariable and Multivariable Joint Frailty Models

Each covariate was first evaluated using univariable joint gamma frailty models for readmission and death. This strategy allowed assessment of marginal covariate effects on both processes while accounting for unobserved heterogeneity and their mutual dependence. Table 1 summarizes the estimated hazard ratios and 95% confidence intervals for the univariable analyses.

Table 1: Univariable joint gamma frailty models for readmission and death

	Readmission		Death	
Variable	HR (95% CI)	p-value	HR (95% CI)	p-value
Complicated atrial fibrillation (CAF)	2.28 (1.65–3.16)	< 0.001	11.93 (4.15–34.30)	< 0.001
Acute respiratory failure (ARF)	8.15 (2.66–24.93)	< 0.001	15.62 (0.69–353.1)	0.08
Extracorporeal circulation (ECC)	2.98 (2.12–4.17)	< 0.001	8.76 (3.93–19.56)	< 0.001
Urea (standardized)	1.37 (1.21–1.56)	< 0.001	3.10 (1.92–5.03)	< 0.001
Hematocrit (per 10 units)	0.51 (0.39–0.66)	< 0.001	0.20 (0.07–0.50)	0.0018
Creatinine (log-transformed)	2.14 (1.58–2.90)	< 0.001	35.75 (19.64–65.08)	< 0.001

In single-covariate joint frailty models, the estimated frailty variance $\hat{\theta}$ ranged between 0.91 and 0.93, with all estimates highly significant. The association parameter α, linking the readmission and death processes, ranged between 4.94 and 5.72, indicating a strong positive association. These results suggest moderate unobserved heterogeneity among patients and a substantial dependence between readmission and death.

Variables showing statistical significance in univariable analyses or deemed clinically relevant were subsequently considered for inclusion in a multivariable joint frailty model. Model selection was based on penalized marginal log-likelihood and likelihood cross-validation criteria, complemented by AIC and BIC. The final multivariable model is presented in Table 2.

Table 2: Multivariable joint gamma frailty model for readmission and death

	Readmission		Death	
Variable	HR (95% CI)	p-value	HR (95% CI)	p-value
Complicated atrial fibrillation (CAF)	1.96 (1.38–2.79)	< 0.001	3.71 (1.44–9.58)	< 0.001
Extracorporeal circulation (ECC)	2.47 (1.74–3.53)	< 0.001	4.32 (1.78–10.48)	0.001
Urea (standardized)	1.20 (1.05–1.37)	0.008	2.39 (1.59–3.59)	< 0.001
Hematocrit (per 10 units)	0.61 (0.43–0.88)	0.008	0.45 (0.25–0.81)	0.007

In the final multivariable model, the frailty variance was estimated at $\hat{\theta} = 0.91$ (standard error 0.06, $p < 0.001$), confirming the presence of moderate unobserved heterogeneity among patients. The association parameter was estimated at $\hat{\alpha} = 4.55$ (standard error 0.78, $p < 0.001$), indicating a strong and stable dependence between readmission and death. We observe that all the variables studied are risk factors for readmission and death, with the exception of hematocrit, which is protective, in line with established clinical understanding.

6.5.2 Model Diagnostics and Interpretation

The proportional hazards assumption for the readmission process was assessed using tests conditional on the frailty term, as implemented in the `frailtyEM` package. No evidence of violation of the proportional hazards assumption was detected at either the individual or global level (Table 3).

Table 3: Tests of the proportional hazards assumption for readmission (conditional on the frailty term)

Variable	Chi-square	df	p-value
Complicated atrial fibrillation (CAF)	1.40	1	0.24
Extracorporeal circulation (ECC)	0.14	1	0.71
Urea (standardized)	2.03	1	0.15
Hematocrit (per 10 units)	0.02	1	0.88
Global test	3.16	4	0.53

Figure 2 displays the estimated baseline survival functions for readmission and death adjusted for unobserved heterogeneity. The population-level survival curves reflect early failures among more vulnerable patients, illustrating the impact of frailty on the marginal survival experience.

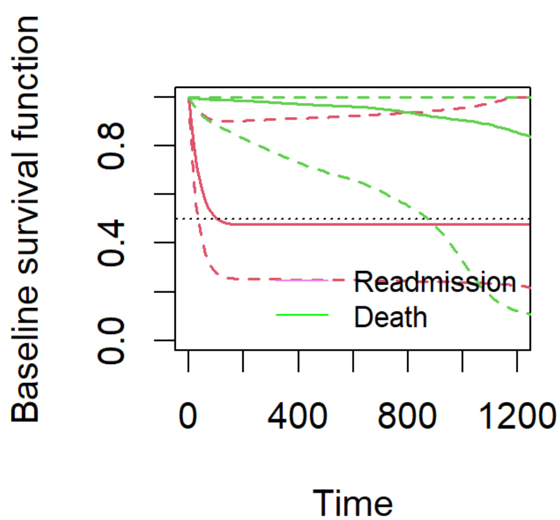

Figure 2: Baseline survival functions for readmission and death adjusted for unobserved heterogeneity. Dotted lines indicate 95% confidence bands.

Figure 3 shows the corresponding baseline hazard functions. The hazard of readmission is markedly elevated shortly after discharge and rapidly decreases over time, whereas the hazard of death remains low during early follow-up and increases at later times. These distinct temporal patterns underscore the relevance of joint frailty modeling for capturing the dynamics of non-terminal and terminal events.

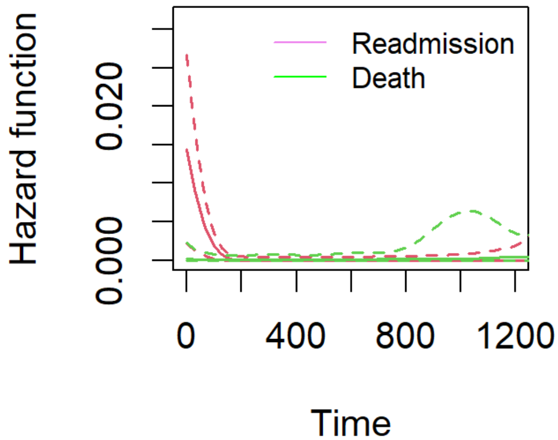

Figure 3: Baseline hazard functions for readmission and death adjusted for unobserved heterogeneity. Dotted lines represent 95% confidence bands.

Overall, the joint frailty approach provided a coherent and interpretable framework for analyzing readmission and death simultaneously. Results were stable across different model specifications and consistent with clinical expectations, despite inherent limitations of the dataset, including a limited number of terminal events and potential multicollinearity among covariates.

7 Discussion and Conclusions

This paper provides an integrated methodological perspective on frailty models for survival analysis, with particular emphasis on recurrent events, semi-competing risks, and joint frailty formulations. Rather than focusing on a specific clinical application, the primary contribution lies in clarifying how unobserved heterogeneity and dependence between event processes fundamentally affect inference, interpretation, and model adequacy in realistic survival settings.

From a methodological standpoint,the paper highlights how frailty models extend classical proportional hazards formulations by explicitly accounting for latent heterogeneity that cannot be captured by observed covariates alone. The distinction between conditional (frailty-specific) and marginal (population-level) quantities is central: while proportional hazards assumptions may hold conditionally, marginal hazards and survival functions typically exhibit non-proportional behavior due to selection effects induced by frailty. Ignoring this distinction may lead to misleading conclusions, particularly when interest lies in population-level risk patterns.

The illustrative example highlights the relevance of joint frailty models in settings where recurrent non-terminal events occur in the presence of a terminal event. Separate analyses of readmission and death obscure their dependence structure and fail to account for shared latent susceptibility. By linking both processes through a common frailty term and an association parameter, joint frailty models provide a coherent framework for semi-competing risks data and allow meaningful interpretation of both within-process and cross-process effects.

An important practical message concerns the role of distributional assumptions for the frailty term. While several frailty distributions are theoretically appealing, their computational properties and identifiability differ substantially in finite samples. In the illustrative analysis, the Gamma frailty distribution consistently yielded stable estimates and superior model fit according to likelihood-based criteria. More flexible alternatives may offer conceptual advantages, but often at the cost of increased numerical instability or reliance on approximation methods.

The comparative assessment of estimation strategies and software implementations underscores that methodological validity does not depend solely on the formal model specification. Different software packages implement distinct likelihood formulations, data structures, and numerical algorithms, which may lead to convergence issues, near-singular Hessian matrices, or boundary estimates of variance and association parameters. Such numerical behavior should not be interpreted automatically as model failure, but rather as a signal of weak identifiability, limited information on latent components, or flat likelihood surfaces inherent to complex models.

Several take-home methodological messages emerge from this work:

- When recurrent and terminal events are jointly observed, modeling them separately may result in biased or incomplete inference; joint frailty models offer a principled alternative.
- Unobserved heterogeneity alters marginal survival and hazard behavior even when conditional proportional hazards assumptions hold.
- The choice of frailty distribution involves a trade-off between theoretical flexibility and numerical stability, particularly in finite samples.
- Software implementation details are an integral part of the modeling process and must be carefully aligned with the underlying scientific and statistical assumptions.

Despite its methodological focus, this study has limitations. The illustrative dataset contains a relatively small number of terminal events relative to the number of covariates, which may affect precision of association parameter estimates. In addition, attention was restricted to parametric frailty distributions and proportional hazards formulations, which, although widely used, may not fully capture all aspects of time-varying effects or complex dependence structures.

Alternative approaches to semi-competing risks data, including marginal structural illness–death models, have also been proposed in recent literature (Zhang et al., 2023).

Future research directions include deeper integration of frailty and multi-state modeling frameworks. Bridging these perspectives would allow for simultaneous modeling of complex transition dynamics and unobserved heterogeneity, providing more comprehensive tools for the analysis of longitudinal and survival data

Further methodological work is also needed to improve numerical stability and inference for variance and association parameters in joint frailty models, particularly in moderate sample sizes.

In conclusion, frailty models constitute a powerful and conceptually coherent class of methods for analyzing survival data with unobserved heterogeneity and dependent event processes. When applied with careful attention to assumptions, identifiability, and implementation, they provide insights that are not accessible through standard survival models and contribute meaningfully to methodological practice in applied survival analysis.

References

Balan, T. A. and Putter, H. (2019). frailtyEM: An R package for estimating semiparametric shared frailty models. *Journal of Statistical Software*, 90(7):1–29.

Breslow, N. (1972). Discussion of regression models and life-tables by D. R. Cox. *Journal of the Royal Statistical Society: Series B*, 34(2):216–217.

Cox, D. R. (1972). Regression models and life-tables(with discussion). *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–220.

Duchateau, L. and Janssen, P. (2010). *The Frailty Model*. Statistics for Biology and Health. Springer, New York, 2 edition.

Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika*, 73(3):671–678.

Hougaard, P. (1999). Multi-state models: A review. *Lifetime Data Analysis*, 5:239–264.

Hougaard, P. (2012). *Analysis of Multivariate Survival Data*. Springer, New York, 2 edition.

Leão, J., Leiva, V., Saulo, H., and Tomazella, V. (2018). A survival model with Birnbaum–Saunders frailty for uncensored and censored cancer data. *Brazilian Journal of Probability and Statistics*, 32(4):707–729.

Monaco, J. V., Gorfine, M., and Hsu, L. (2018). General semiparametric shared frailty model: Estimation and simulation with frailtySurv. *Journal of Statistical Software*, 86(4):1–42.

Putter, H., Fiocco, M., and Geskus, R. B. (2007). Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine*, 26:2389–2430.

R Core Team (2025). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rondeau, V., Mazroui, Y., and Gonzalez, J. (2012). frailtypack: An R package for the analysis of correlated survival data with frailty models using penalized likelihood estimation or parametrical estimation. *Journal of Statistical Software*, 47(4):1–28.

Therneau, T. (2023). *survival: A package for survival analysis in R*. R package version 3.5.5.

Therneau, T. (2024). *coxme: Mixed effects Cox models*. R package version 2.2-22.

Wienke, A. (2010). *Frailty Models in Survival Analysis*. Chapman and Hall/CRC, New York.

Wu, T., Feng, C., and Li, L. (2023). A comparison of parameter estimation methods for shared frailty models. *arXiv preprint*.

Zhang, Y., Ying, L., Edland, S., White, L., and Xu, R. (2023). Marginal structural illness–death models for semi-competing risks data. *arXiv preprint*.