Smooth Transformation Models for Survival Analysis: A Tutorial Using R

Sandra Siegfried
Universität Zürich

Bálint Tamási
Universität Zürich

Torsten Hothorn
Universität Zürich

Abstract

Over the last five decades, we have seen strong methodological advances in survival analysis, using parametric methods and, more prominently, methods based on non-/semi-parametric estimation. As the methodological landscape continues to evolve, the task of navigating through the multitude of methods and identifying available software resources is becoming increasingly challenging – especially in more complex scenarios, such as when dealing with interval-censored or clustered survival data, non-proportional hazards, or dependent censoring.

This tutorial explores the potential of using the framework of smooth transformation models for survival analysis in the R system for statistical computing. This framework provides a unified maximum-likelihood approach that covers a wide range of survival models, including well-established ones such as the Weibull model and a fully parametric version of the famous Cox proportional hazards model, and various extensions for more complex scenarios. We explore models for non-proportional/crossing hazards, dependent censoring, clustered observations and extensions towards personalised medicine within this framework.

Using survival data from a two-arm randomised controlled trial on rectal cancer therapy, we demonstrate how survival analysis tasks can be seamlessly navigated in R within this framework using the implementation provided by the **tram** package, and few related packages.

Keywords: non-proportional hazards, dependent censoring, clustered observations, personalised medicine, survival trees, R.

1. Introduction

In "parametric" survival analysis, researchers commonly rely on the Weibull model or alternative accelerated failure time models. To achieve greater flexibility and overcome the strict distributional assumptions underlying these models, researchers often need to turn to non-/semi-parametric methods to analyse their survival data. When it comes to non-/semi-parametric approaches, however, overcoming issues tied to interval-censored or truncated observations can prove challenging due to their limited availability in standard software im-

Please cite this work as: Sandra Siegfried, Bálint Tamási and Torsten Hothorn (2026), Smooth Transformation Models for Survival Analysis: A Tutorial Using R, *Statistical Methods in Medical Research*, 35(5), 980–997 https://doi.org/10.1177/09622802251414595.

plementations.

Furthermore, when aiming to fit models that handle crossing or non-proportional hazards, clustered observations, or dependent censoring, researchers often find themselves navigating a complex landscape of diverse software implementations. Even the same models can be difficult to compare across different implementations, because different parametrisations, estimation strategies or optimisation procedures are used.

This becomes even more challenging when comparing different models computed from different packages – emphasising the benefit of a unified framework that facilitates seamless transitions between different models and is based on a common core infrastructure for model parametrisation, estimation strategy, and optimisation procedure.

To tackle these challenges, researchers may explore the potential of the **tram** add-on package (Hothorn *et al.* 2025) in R (**?**), which offers a flexible framework for survival analysis. The **tram** package implements a user-friendly interface to smooth transformation models (Hothorn *et al.* 2018), which encompass a range of classical survival models including accelerated failure models and the Cox proportional hazards model, as well as many useful extensions and novel model formulations. The package can be easily installed from the Comprehensive R Archive Network (CRAN):

```
R> install.packages("tram")
R> library("tram")
```

All models are implemented in a fully parametric fashion, allowing for straightforward maximum likelihood inference. The estimation of these models is facilitated by the core infrastructure package **mlt** (Hothorn 2020a, 2025a) which provides a unified maximum likelihood framework for general transformation models (Hothorn *et al.* 2018). We further leverage two add-on packages from this family of packages for transformation modelling: The **tramME** package (Tamási 2025, 2024) implementing mixed-effects and non-linear additive effects for smooth transformation models, and the **trtf** package (Hothorn 2025c) for estimating tree-based survival models.

In this tutorial, we will explore a variety of models commonly utilised in survival analysis. The focus of this tutorial lies on the practical implementation and interpretation of these models within the framework of smooth transformation models, rather than on theoretical aspects. Our objective is to provide users with a practical understanding of how to apply these models using R. Through an application to data from a randomised trial on rectal cancer therapy, we showcase how users can seamlessly conduct their survival analysis tasks without the need to navigate through a multitude of packages in R.

In Section 3, we discuss models for independent observations. We start with the well-known Weibull model, and then, to introduce more flexibility and overcome the log-linear log-cumulative hazard assumption inherent to the Weibull model, we explore a fully parametric version of the Cox model. We further discuss the estimation of stratified log-cumulative hazard functions to account for baseline risk variations across patient strata. Moving beyond the assumption of proportional hazards, we showcase models that challenge this assumption. We discuss a location-scale version of the Cox model, accommodating scenarios with non-proportional/crossing hazards, and models estimating time-varying treatment effects.

Addressing scenarios where the assumption of independent censoring might not be justified, we discuss a copula proportional hazards model, that accommodates dependent censoring

(Section 4). For clustered observations we employ mixed-effects Cox models and alternative models featuring marginally interpretable hazard ratios in Section 5. Our tutorial also explores the domain of personalised medicine, presenting models that incorporate covariate-dependent treatment effects and survival trees (Section 6). In Section 7, we explore further extensions, including topics like frailty models, model estimation using the non-parametric likelihood, covariate adjustment and the potential of using these models for sample size estimation of new trials.

This tutorial is composed of the main text, which introduces the models and very briefly shows how to estimate them using smooth transformation models in R. In addition, we present head-to-head comparisons of user-interfaces and numerical results obtained from alternative packages available in the R universe in Supplementary Material A. Both parts come with much more detailed R code for exploring fitted models (for example, plotting model terms, computing confidence intervals, or performing tests), which can be explored in the corresponding demo:

```
R> demo("survtram", package = "tram")
```

In our Supplementary Material A, we conduct a thorough comparison of a subset of the models discussed here which can be estimated using alternative implementations (in total 13 established CRAN packages) and corresponding results obtained with smooth transformation models from **tram** and **tramME**. This quality assurance task not only helped to validate the implementation in **tram** and **tramME** but also led to the identification of problematic special cases and, in some instances, practically relevant discrepancies between different package implementations of the very same model. Moreover, Supplementary Material A presents the different user-interfaces of the different packages side-by-side, such that it becomes simpler to estimate and compare relatively complex models across independent implementations. For the analysis of future survival trials, an assessment of the agreement of such estimates, standard errors, and possibly other model aspects can help to increase trust in reported numerical results and conclusions derived therefrom.

2. Application

In our tutorial, we will work with data from the CAO/ARO/AIO-04 two-arm randomised controlled trial (Rödel *et al.* 2015), a phase 3 study that aimed to compare an experimental regimen with the previously established treatment regimen (control) for locally advanced rectal cancer. In this experimental regimen, Oxaliplatin was added to the control treatment of preoperative Fluorouracil-based chemoradiotherapy and postoperative chemotherapy of locally advanced rectal cancer.

The trial was conducted across 88 centers and included a cohort of 1'236 patients. The patients were randomly allocated to the two treatment arms $W \in \{0, 1\}$, receiving the experimental treatment of Oxaliplatin added to Fluorouracil-based preoperative chemoradiotherapy and postoperative chemotherapy (5-FU + Oxaliplatin, $W = 1$) or the control treatment using Fluorouracil only (5-FU, $W = 0$). Treatment allocation was performed using block-randomisation stratified by study center $j = 1, \ldots, 88$ and the stratum s, which is defined by four categories consisting of a combination of clinical N category, *i.e.,* lymph node involvement (cN0 vs cN+), and clinical T category *i.e.,* tumor grading (cT1-3 vs cT4). The distribution of patients in the two treatment arms across strata is shown in Table 1.

			5-FU	5-FU + Oxaliplatin
cT1-3	:	cN0	163	156
cT4	:	cN0	8	7
cT1-3	:	cN+	411	417
cT4	:	cN+	41	33
Total			623	613

Table 1: Number of patients in each treatment arm stratified by the combination of clinical N and T category.

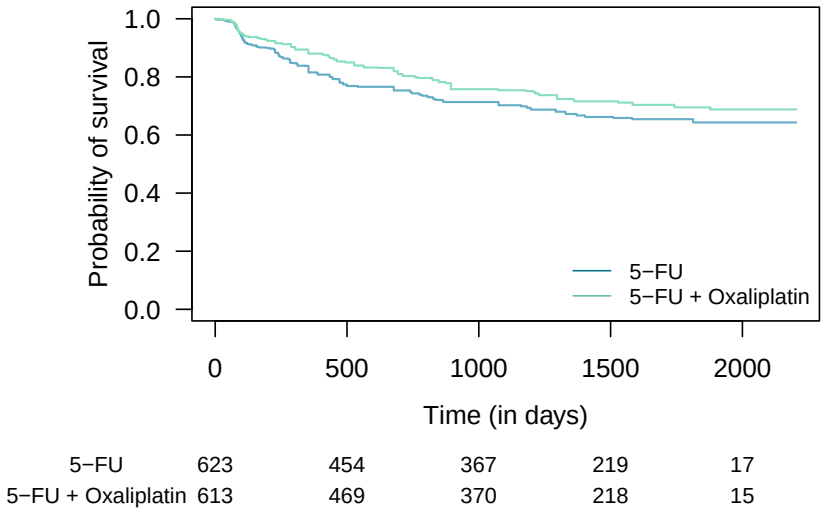


Figure 1: Disease-free survival. The survivor functions of the two treatment arms estimated by the non-parametric Turnbull method are shown together with the number at risk table.

The primary endpoint is disease-free survival, defined as the time $T \in \mathbb{R}^+$ between randomisation and non-radical surgery of the primary tumor (R2 resection), loco-regional recurrence after R0/1 resection, metastatic disease or progression, or death from any cause – whichever occurred first. The observed times encompass a mix of exact observations t for time to death or incomplete removal of the primary tumor, interval-censored observations $t \in (\underline{t}, \bar{t}]$ for the time span from the previous follow-up $\underline{t}$ to the time-point of detecting local or distant metastases $\bar{t}$, and right-censored observations $t \in (t, \infty)$ corresponding to the end of the follow-up period or instances where patients were lost to follow-up. The survivor curves of the primary endpoint (disease-free survival) estimated by the non-parametric Turnbull estimator (Turnbull 1976) are shown for the two treatment arms in Figure 1.

A secondary endpoint considered in the study is overall survival, defined as the time $T \in \mathbb{R}^+$ from randomisation to death from any cause. Notably, all observations t for this endpoint are exact or right-censored. The corresponding survivor curves, estimated non-parametrically by the Kaplan-Meier method (Kaplan and Meier 1958), are shown in Figure 2.

The primary data analysis for this trial was performed by Rödel *et al.* (2015). In their analysis, the treatment effect comparing the effect of the experimental treatment to the effect of the control treatment on disease-free and overall survival was assessed by adjusted log-rank tests and mixed-effects Cox models (both adjusting for the stratified randomisation process),

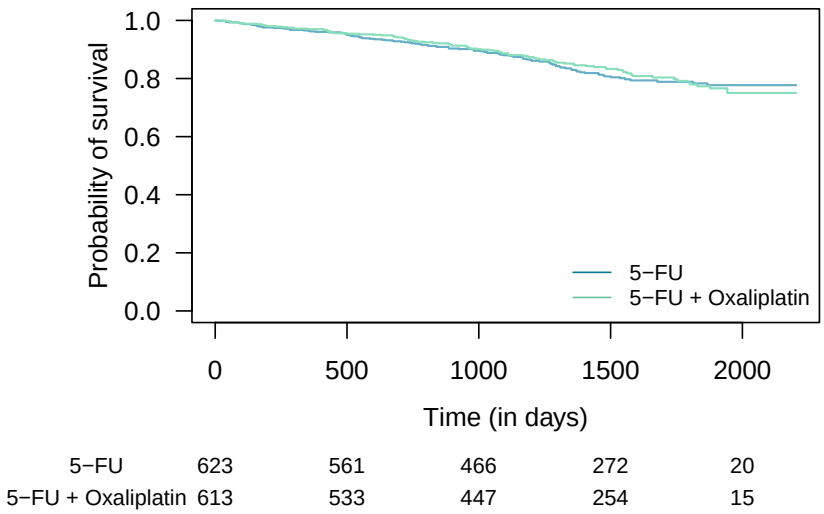


Figure 2: Overall survival. The survivor functions of the two treatment arms estimated by the non-parametric Kaplan-Meier method are shown together with the number at risk table.

treating both survival endpoints as exact. Following, we demonstrate the process of fitting a fully parametric mixed-effects Cox model that accounts for interval-censored event times of the primary endpoint in Section 5, also in terms of a model featuring marginal interpretation of the estimated hazard ratio.

A subsequent post hoc analysis was carried out by Hofheinz *et al.* (2018). In this analysis, the possibility of an age-varying treatment effect on both the primary endpoint of disease-free survival and the secondary endpoint of overall survival was investigated. In Section 6, we demonstrate how such an analysis can be performed within the discussed framework, taking into account interval-censoring. We illustrate two approaches for estimating age-varying effects, using smooth age-varying hazard ratios and age-structured survival trees.

While the analyses conducted by Rödel *et al.* (2015) and Hofheinz *et al.* (2018) made serious efforts to address the research questions effectively, they were limited by the lack of software resources capable of adequately handling interval-censored and correlated observations for the analysis of the primary endpoint. Notably, the first R add-on package capable of estimating Cox models in the presence of interval-censoring was published in 2014 (**coxinterval** package Boruvka and Cook 2014). At the time of the statistical analysis of the primary endpoint, it was impossible to fit mixed-effects models with flexible baseline hazards to interval-censored outcomes. This obstacle was one of the main motivation to develop a comprehensive software package implementing a general class of transformation models with applications in the domain of survival analysis. The corresponding framework implementing smooth transformation models (Hothorn *et al.* 2018) in R helps to address such and other practically relevant limitations. In this tutorial, we present analyses that the CAO/ARO/AIO-04 study investigators would have liked to have been able to perform a decade ago.

3. Independent observations

3.1. Survival models

Weibull proportional hazards model

Probably one of the most renowned parametric model in survival analysis is the Weibull model (Wei 1992), where the response T conditional on treatment assignment $W = w$ is assumed to follow a Weibull distribution. We consider the Weibull proportional hazards model with survivor functions conditional on the treatment assignment parametrised as,

$$\begin{aligned} S_0(t) &= \mathbb{P}(T > t \mid W = 0) = \exp\left\{-\exp\left[\vartheta_1 + \vartheta_2 \log(t)\right]\right\}, \qquad \vartheta_2 > 0, \\ S_1(t) &= \mathbb{P}(T > t \mid W = 1) = \exp\left\{-\exp\left[\vartheta_1 + \vartheta_2 \log(t) - \beta\right]\right\}, \end{aligned}$$

with the general formula,

$$S_w(t) = \mathbb{P}(T > t \mid W = w) = \exp\left\{-\exp\left[\vartheta_1 + \vartheta_2 \log(t) - \beta w\right]\right\}. \tag{1}$$

The log-cumulative baseline hazard $\log(-\log(S_0(t))) = \log(\Lambda_0(t))$ here is given by $h(t) = \vartheta_1 + \vartheta_2 \log(t)$, assuming a linear shift β on the scale of log-time $\log(t)$. The model not only assumes proportional hazards, with hazard ratio $\Lambda_1(t)/\Lambda_0(t) = \lambda_1(t)/\lambda_0(t) = \exp(-\beta)$, but is also an accelerated failure time model

$$\begin{aligned} \log(T) &= \frac{-\vartheta_1 + \beta w + Z}{\vartheta_2}, \qquad Z \sim \text{MEV} \\ &= -\frac{\vartheta_1}{\vartheta_2} + \frac{\beta}{\vartheta_2} w + \frac{1}{\vartheta_2} Z = \alpha + \tilde{\beta} w + \sigma Z, \end{aligned}$$

with the errors Z following a minimum extreme value distribution (MEV). Consequently, $T \mid W = w$ follows a Weibull distribution (Kalbfleisch and Prentice 2002, Chapter 2) with intercept $\alpha = -\vartheta_1 \vartheta_2^{-1}$, scale parameter $\sigma = \vartheta_2^{-1}$ and log-acceleration factor $\tilde{\beta} = \beta \vartheta_2^{-1}$. The model implies that, for the experimental arm time T is accelerated by $\exp(\tilde{\beta})$, that is $T_1 = \exp(\tilde{\beta}) T_0$, thus the probability of disease-free survival for the experimental arm is given by $S_1(t) = S_0(\exp(-\tilde{\beta})t)$.

Alternatively, different distributions for Z, such as the normal or logistic distribution, can be specified, leading to the formulation of log-normal or log-logistic models.

Parameter estimation of the Weibull model is straightforward using maximum likelihood, because the distribution function can be directly evaluated and thus allows to effectively handle interval-censored or truncated data, as will be discussed in Section 3.2.

Flexible proportional hazards model

The assumption of a log-linear log-cumulative baseline hazard function h, implied by the Weibull model, can be relaxed by replacing $\log(\Lambda_0(t)) = h(t) = \vartheta_1 + \vartheta_2 \log(t)$ with a more flexible function $h(t) = \boldsymbol{a}(t)^\top \boldsymbol{\vartheta}$ defined in terms of smooth spline basis functions $\boldsymbol{a}$ and corresponding parameters $\boldsymbol{\vartheta}$. This yields the following model

$$S_w(t) = \mathbb{P}(T > t \mid W = w) = \exp\left\{-\exp\left[h(t) + \beta w\right]\right\}, \tag{2}$$

where the hazard ratio is given by $\Lambda_1(t)/\Lambda_0(t) = \lambda_1(t)/\lambda_0(t) = \exp(\beta)$. McLain and Ghosh (2013) and Hothorn *et al.* (2018) proposed parametrising $h(t) = \boldsymbol{a}(t)^\top \boldsymbol{\vartheta}$ as polynomials in

Bernstein form $h(t) = \boldsymbol{a}_{\text{Bs},P-1}(t)^\top \boldsymbol{\vartheta}$ of order $P-1$. The corresponding model (2) is a fully parametric version of the otherwise semi-parametric Cox proportional hazards model (Cox 1972). The latter treats h as an infinite dimensional nuisance parameter which is profiled out from the likelihood. This leads to the partial likelihood, through which the log-hazard ratio β can be estimated (Cox 1975). In contrast, all parameters of model (2) are estimated simultaneously by maximum likelihood. The parametrisation of h in terms of basis functions and corresponding parameters allows to specify of a flexible, yet fully parametric, monotonically increasing log-cumulative hazard function. This is achieved under appropriate constraints $\vartheta_p \leq \vartheta_{p+1}$ for $p \in 1, \ldots, P-1$ (Hothorn *et al.* 2018). Adopting this specific parametrisation for the log-cumulative baseline hazard function $\log(\Lambda_0(t)) = h(t)$ facilitates the computation of the corresponding density $f_0(t)$ and distribution function $F_0(t)$, thus allowing for straightforward parameter estimation using maximum likelihood. This holds true even when dealing with interval-censored or truncated observations.

Within the **tram** add-on package, the order, $P-1$, of polynomials in Bernstein form is not determined in a data-driven way. The default $P-1=6$ is typically a good compromise between flexibility of $h(t)$ and computing time needed to optimise the log-likelihood. Fixed P also facilitates standard maximum likelihood inference. Because of the monotonicity constraint on h, the transformation function h it not prone to overfit and thus, in principle, P can be chosen much larger. The effect of larger P on estimates of other model parameters and their standard errors is very small (see, for example, the log-hazard ratios in Figure 5 provided by Hothorn (2020a) and the empirical comparison to non-parametric models Tian *et al.* (2020)). However, if one is interested in expressions involving the derivative $h'(t)$, which itself is in Bernstein form, the order $P-1$ must be chosen in a data-driven way, for example for density estimation. Sieve maximum likelihood procedures have been suggested in this context, for example in Cox models with log-cumulative baseline hazard functions in Bernstein form McLain and Ghosh (2013).

Stratified proportional hazards model

Accounting for variations in baseline risks among different patient strata identified by variable s, one can employ stratified models that incorporate stratum-specific log-cumulative baseline hazard functions $h(t \mid s)$. These models can be defined by

$$S_w(t \mid s) \quad = \quad \mathbb{P}(T > t \mid S = s, W = w) = \exp\left\{-\exp\left[h(t \mid s) + \beta w\right]\right\},$$

with $h(t \mid s) = \boldsymbol{a}(t)^\top \boldsymbol{\vartheta}(s)$ and global hazard ratio $\frac{\Lambda_1(t|s)}{\Lambda_0(t|s)} = \frac{\lambda_1(t|s)}{\lambda_0(t|s)} = \exp(\beta)$, assuming that the treatment effect is the same across all patient strata s.

One could further relax the restriction of a global treatment effect, allowing for an interaction of the treatment assignment $W = w$ and the stratum s by formulating the log-cumulative hazard as $h(t \mid s) + w\beta + \boldsymbol{\gamma}^\top (w \times s)$.

Non-proportional hazards model

Extending beyond the proportional hazards assumption, an additional treatment-dependent model term can be estimated. Burke and MacKenzie (2017) introduced the multi-parameter extension to the Weibull model (1) in the context of survival analysis, specifically outlining its use for interval-censored observations (Peng *et al.* 2020).

A similar extension can be made to the more flexible, parametric, Cox model (2), by additionally estimating a scale term γ for the experimental arm (Siegfried *et al.* 2023),

$$S_w(t) \quad = \quad \mathbb{P}(T > t \mid W = w) = \exp\left\{-\exp\left[\sqrt{\exp(\gamma w)}\, h(t) + \beta w\right]\right\}.$$

In this case the ratio of the cumulative hazards, $\Lambda_1(t)/\Lambda_0(t)$, is a non-proportional function of t. The corresponding bivariate score test (Section 3.1.2 of Siegfried *et al.* 2023) further allows to test the null hypothesis of equal survival, *i.e.,* $\beta = \gamma = 0$, without relying on the assumption of proportional hazards.

Time-varying hazards model

Accounting for changing effects of the treatment over time, we can further extend beyond the proportional hazards assumption and estimate a model incorporating a time-varying treatment effect,

$$S_w(t) \quad = \quad \mathbb{P}(T > t \mid W = w) = \exp\left\{-\exp\left[h(t) + \beta(t) w\right]\right\}.$$

Here, the model introduces a time-varying shift $\beta(t)$ in the log-cumulative hazard function $\log(\Lambda_1(t)) = \log(\Lambda_0(t)) + \beta(t)$, thereby relaxing the assumption of a constant log-hazard ratio β. The shift $\beta(t)$ is also parameterised in terms of a polynomial in Bernstein form, thus allowing to estimate a time-varying ratio of the cumulative hazards $\Lambda_1(t)/\Lambda_0(t) = \exp(\beta(t))$.

3.2. Likelihood

From the above models, the log-likelihoods for exact or independently right-, left- or interval-censored and truncated observations are easily deducible. We here consider the most general case where the log-cumulative hazard function is given by $h(t \mid w, s) = \sqrt{\exp(\gamma w)}\, h(t \mid s) + \beta w$.

For exact continuous observations (t, w, s), the corresponding likelihood contributions are given by

$$\ell(\boldsymbol{\vartheta}(s), \beta, \gamma \mid T = t) = \log\left\{f\left[h(t \mid w, s)\right]\right\} + \log\left\{h'(t \mid w, s)\right\},$$

with the standard minimum extreme value density $f(z) = \exp(z - \exp(z))$ and the derivative of the log-cumulative hazard function with respect to t, $h'(t \mid w, s) = \sqrt{\exp(\gamma w)}\, \boldsymbol{a}'(t)^\top \boldsymbol{\vartheta}(s)$.

Because the transformation function h, defining the log-cumulative baseline hazard function, is parametrised in terms of polynomials in Bernstein form, where the basis functions $\boldsymbol{a}_{\text{Bs},P-1}(t) \in \mathbb{R}^P$ are specific beta densities (Farouki 2012), it is straightforward to obtain the derivatives of the basis functions with respect to t, leading to $h'(t \mid s) = \boldsymbol{a}'_{\text{Bs},P-1}(t)^\top \boldsymbol{\vartheta}(s)$.

Under independent left-, right- or interval-censored event times $(\underline{t}, \bar{t}]$ the exact log-likelihood contribution is

$$\ell(\boldsymbol{\vartheta}(s), \beta, \gamma \mid T \in (\underline{t}, \bar{t}]) = \log\left\{\mathbb{P}(T \in (\underline{t}, \bar{t}] \mid w, s)\right\} = \log\left\{S_w(\underline{t} \mid w, s) - S_w(\bar{t} \mid w, s)\right\}.$$

For observations that are right-censored at time t the datum is given by $(\underline{t}, \bar{t}] = (t, \infty)$ and for left-censored observations it is $(\underline{t}, \bar{t}] = (0, t]$.

In cases where event times are subject to random left-, right-, or interval-truncation $(t_\text{l}, t_\text{r}] \subset \mathbb{R}^+$, the above log-likelihood contributions change to

$$\ell(\boldsymbol{\vartheta}(s), \beta, \gamma \mid T \in (\underline{t}, \bar{t}]) - \ell(\boldsymbol{\vartheta}(s), \beta, \gamma \mid T \in (t_\text{l}, t_\text{r}])$$

with $t_\text{l} = 0$ for right-truncated and $t_\text{r} = \infty$ for left-truncated observations. Such considerations are relevant in scenarios involving a late entry approach, for instance, resulting in left-truncated observations, where one is interested in modelling $\mathbb{P}(t > T \mid T \in (t_\text{l}, \infty))$, or for modelling time-varying covariates.

3.3. Application

Now, turning our attention to the CAO/ARO/AIO-04 two-arm randomised controlled trial, we aim to estimate the previously discussed models using the unified maximum likelihood framework provided by the R add-on package **tram** (Hothorn *et al.* 2025). We fit the models to the primary endpoint of disease-free survival T estimating the treatment effect corresponding to the assignment $W =$ `randarm`.

The disease-free survival times are stored as `iDFS`, an object of class '`Surv`', which includes a mix of exact, interval-, and right-censored observations. This '`Surv`' object can be specified with `Surv(CAOsurv$iDFStime, CAOsurv$iDFStime2, type = "interval2")` (Therneau 2024; Therneau and Grambsch 2000). Exact observations are represented by two identical time points, for interval-censored observations, the two times define the period within which the event occurred and right-censored observations are represented by an interval from the last visit to infinity.

We start by fitting the Weibull model (Section 3.1.1) using the `Survreg()` function.

```
R> Survreg(iDFS ~ randarm, data = CAOsurv, dist = "weibull")
```

Coefficient	Estimate	Std. Error	95%-Wald CI
ϑ_1	−6.231	0.265	−6.752 to −5.711
ϑ_2	0.733	0.036	0.663 to 0.803
β	0.229	0.106	0.020 to 0.438

Log-Likelihood	Likelihood Ratio Test	Score Test	Permutation Score Test
−2'281.17	$p = 0.031$	$p = 0.031$	$p = 0.035$

The model quantifies the treatment effect through a hazard ratio $\exp(-\hat{\beta}) = 0.795$, comparing the hazards of the experimental arm to the hazards of the control arm. The results indicate a reduction in hazards for patients receiving the experimental treatment compared to the control treatment. This reduction in hazards translates to a prolonged disease-free survival time in the experimental arm. Since the model is a proportional hazards counterpart of the Weibull accelerated failure time model fitted by `survreg()` from the **survival** package (Therneau and Grambsch 2000; Therneau 2024), the estimate can also be translated into a log-acceleration factor $\hat{\tilde{\beta}} = \hat{\beta}\hat{\vartheta}_2^{-1} = 0.312$. This implies that the disease-free survival time T is prolonged by $\exp(\hat{\tilde{\beta}}) = 1.367$ in the experimental arm, compared to the control arm.

Next, we fit the flexible proportional hazards model (Section 3.1.2) using the `Coxph()` function from the **tram** package.

```
R> Coxph(iDFS ~ randarm, data = CAOsurv)
```

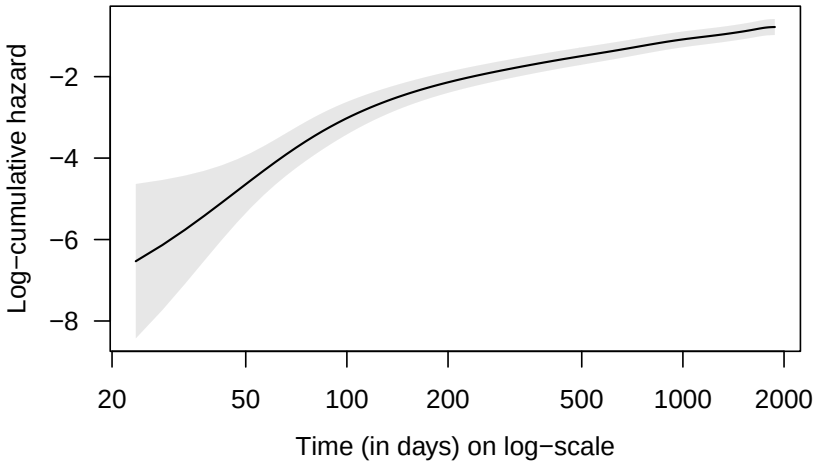


Figure 3: Flexible proportional hazards model. Log-cumulative baseline hazard function and corresponding 95%-confidence band estimated by the model.

Coefficient	Estimate	Std. Error	95%-Score CI
β	-0.231	0.107	-0.439 to -0.022
Log-Likelihood	Likelihood Ratio Test	Score Test	Permutation Score Test
$-2'242.25$	$p = 0.030$	$p = 0.030$	$p = 0.030$

The fitted model is a fully parametric version of the famous Cox model, otherwise estimated semi-parametrically using the partial likelihood (as implemented in the **survival** package in the `coxph()` function). Here the log-cumulative hazard function is specified in terms of polynomials in Bernstein form, by default of order $P - 1 = 6$, specifying the transformation function $h(t) = \boldsymbol{a}_{\text{Bs},6}(t)^\top \boldsymbol{\vartheta}$. The fully parametric approach enables the straightforward incorporation of interval-censored disease-free survival times. Figure 3 illustrates the estimated log-cumulative baseline hazard function $\log(\hat{\Lambda}_0(t)) = \hat{h}(t) = \boldsymbol{a}_{\text{Bs},6}(t)^\top \hat{\boldsymbol{\vartheta}}$ along with the 95%-confidence band, revealing a non-linear function of log-time. The band was obtained from simultaneous confidence intervals for a dense grid of time points utilising the asymptotic normality of the maximum likelihood estimator $\hat{\boldsymbol{\vartheta}}$ and the fact that h was parameterised as a contrast. The estimated hazard ratio is $\exp(\hat{\beta}) = 0.794$, indicating reduced hazards in the experimental arm. The software implementation further allows to compute a corresponding 95%-permutation score confidence interval which ranges from 0.645 to 0.978.

To further accommodate for varying log-cumulative baseline hazard functions $\Lambda_0(t \mid s)$ across patient strata s (here identified by `strat`), we can fit a stratified model (Section 3.1.3).

For the package vignette only, we speed things up a bit allowing the optimiser to stop early:

```
R> Coxph(iDFS | strat ~ randarm, data = CAOsurv,
+    optim = fastopt)
```

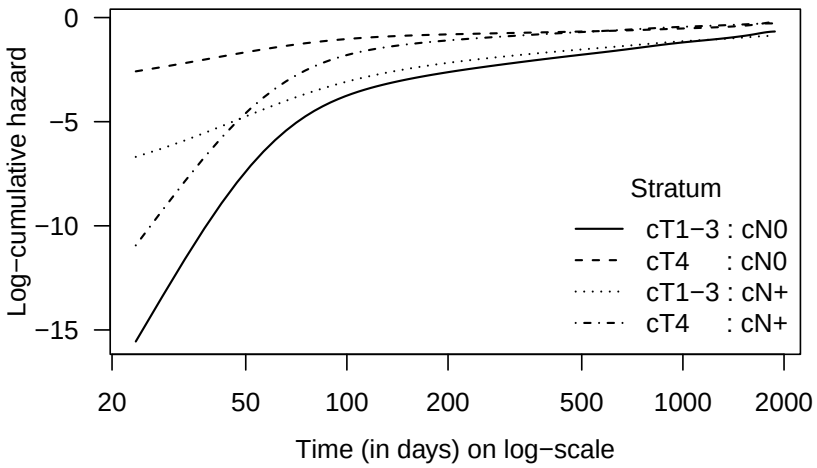


Figure 4: Stratified proportional hazards model. Log-cumulative baseline hazard functions $\Lambda_0(t \mid s)$ estimated by the model are shown separately for each stratum s.

Coefficient	Estimate	Std. Error	95%-Score CI
β	−0.228	0.107	−0.436 to −0.020
Log-Likelihood	**Likelihood Ratio Test**	**Score Test**	**Permutation Score Test**
−2'213.93	$p = 0.031$	$p = 0.032$	$p = 0.033$

The model estimates separate smooth log-cumulative baseline hazard functions for each stratum s, as illustrated in Figure 4, but provides an estimate of the global hazard ratio $\exp(\hat{\beta}) = 0.797$, indicating a reduction of the hazard in the experimental arm by 0.797 relative to the hazard in the control arm across all stratum.

Moving away from the proportional hazards assumption, we can fit a non-proportional hazards model (a location-scale version of the Cox model, Section 3.1.4) using the same function.

```
R> Coxph(iDFS ~ randarm | randarm, data = CAOsurv)
```

Coefficient	Estimate	Std. Error	95%-Wald CI
β	−0.091	0.163	−0.411 to 0.229
γ	0.257	0.203	−0.140 to 0.654
Log-Likelihood	**Likelihood Ratio Test**	**Bivariate Permutation Score Test**	
−2'241.46	$p = 0.043$	$p = 0.027$	

The ratio of the cumulative hazards $\widehat{\Lambda_1}(t)/\widehat{\Lambda_0}(t)$, shown in Figure 5, no longer remains proportional but varies over time. The curve indicates a pronounced initial reduction in cumulative hazards for the experimental arm compared to the control arm, which gradually decreases as time progresses. This suggests that the treatment effect is stronger in the beginning. The corresponding bivariate score test, which tests the null hypothesis of equal survival without assuming proportional hazards, further indicates evidence for non-equal disease-free survival times.

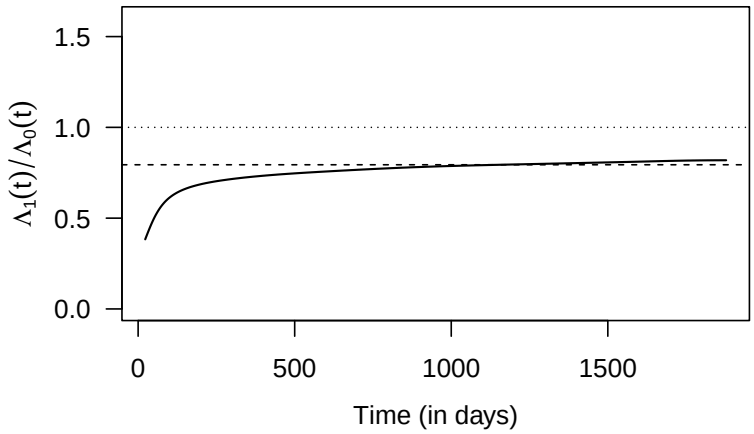


Figure 5: Non-proportional hazards model. $\Lambda_1(t)/\Lambda_0(t)$ (solid line) estimated by the model is shown alongside the constant hazard ratio estimated from the proportional hazards model (dashed line) over time t.

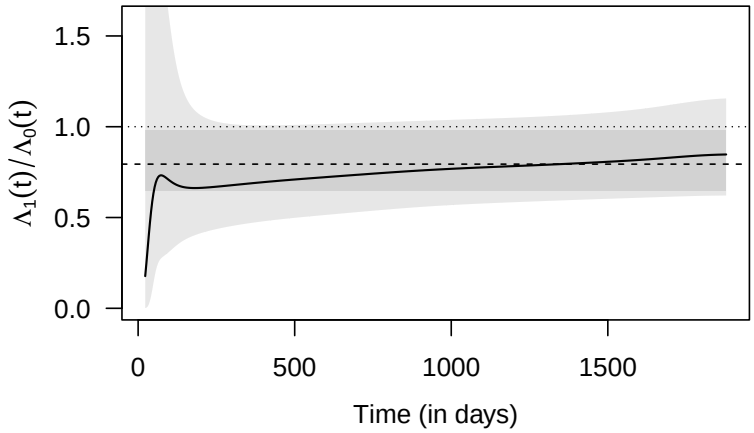


Figure 6: Time-varying hazards model. $\Lambda_1(t)/\Lambda_0(t)$ and corresponding 95%-confidence bands over time t (solid line) estimated by the model is shown alongside the constant hazard ratio estimated from the proportional hazards model (dashed line). The log-likelihood of the model is $-2'240.21$.

Finally, we fit the model featuring a time-varying treatment effect (Section 3.1.5).

```
R> Coxph(iDFS | randarm ~ 1, data = CAOsurv)
```

The treatment effect $\Lambda_1(t)/\Lambda_0(t) = \exp(\beta(t))$ is a function of time, as shown in Figure 6. The curve again demonstrates a reduction in hazards for the experimental arm compared to the control arm, which is more substantial in the beginning and gradually becomes less prominent as time progresses.

4. Dependent censoring

4.1. Copula proportional hazards model

Until now, the models, we have discussed, have been constructed under the assumption of independent censoring, which implies that the censoring times C are independent of the event times T given the treatment assignment $W = w$, that is $T \perp\!\!\!\perp C \mid W = w$. We can however move beyond relying on this assumption and allow the model to capture potential dependence between the censoring times C and event times T.

We explore models discussed in recent work of Czado and Van Keilegom (2023), which involve a joint model for T and C employing a bivariate Gaussian copula of the marginal survivor functions $S_w(t)$ and $G_w(c)$ of T and C respectively,

$$\mathbb{P}(T \leq t, C \leq c \mid W = w) = \Phi_{0,\boldsymbol{R}(\xi)} \left\{ \Phi^{-1}\left[1 - S_w(t)\right], \Phi^{-1}\left[1 - G_w(c)\right] \right\}$$

with correlation matrix

$$\boldsymbol{R}(\xi) = \begin{bmatrix} 1 & -\xi/\sqrt{1+\xi^2} \\ -\xi/\sqrt{1+\xi^2} & 1 \end{bmatrix}, \qquad \xi \in (-\infty, \infty)$$

to account for the association between T and C. Deresa and Van Keilegom (2024) recently demonstrated that the above model maintains identifiability when the marginal survivor functions $S_w(t)$ and $G_w(c)$ are described by a flexible proportional hazards model (2) and a model that assumes a Weibull distribution (1), respectively. This allows to estimate the dependence parameter ξ and other model parameters from the observed data. A dependence parameter ξ of zero corresponds to uncorrelated event times T and censoring times C, thus indicating lack of evidence for dependent censoring.

4.2. Application

Returning to our application, where we previously assumed independent censoring of disease-free survival times, we now aim to address the potential scenario where loss of follow-up time $C \in \mathbb{R}^+$ in the two arms is not independent of the overall survival time $T \in \mathbb{R}^+$ (secondary endpoint).

The observed times can be categorised into the following event types (specified in `DepCevent`): The event of interest (corresponding to overall survival), loss of follow-up (dependent censoring), and end of follow-up period (administrative/independent censoring).

	5-FU	5-FU + Oxaliplatin
Administrative censoring	469	466
Event of interest	106	96
Loss of follow-up	48	51

The model accommodating dependent censoring can also be fitted using the `Coxph()` function by appropriately specifying the event in the '`Surv`' object. Again, we stop the optimiser early in order to save time for the vignette compilation

```
R> Coxph(Surv(OStime, event = DepCevent) ~ randarm, data = CAOsurv,
+    optim = fastoptH)
```

Coefficient	Estimate	Std. Error	95%-Wald CI
β_T	−0.032	0.143	−0.312 to 0.248
ξ	0.021	0.454	−0.869 to 0.912
Log-Likelihood			
−3'061.65			

The joint model is fitted based on a Gaussian copula, estimating a marginal flexible proportional hazards model (2) for the overall survival time T and a marginal Weibull proportional hazards model (1) for the loss of follow-up time C, while accounting for independent right-censoring at the end of the follow-up period.

The estimated hazard ratio assessing the treatment effect on overall survival, is $\exp(\hat{\beta}_T) = 0.969$ with a 95%-confidence interval from 0.732 to 1.281. This indicates no evidence for prolonged overall survival in the experimental compared to the control arm. The estimated dependence parameter is $\hat{\xi} = 0.021$, corresponding to a Kendall's $\hat{\tau} = -0.014$. The corresponding 95%-confidence interval from −0.869 to 0.912 for ξ does include zero, providing no evidence for a dependence between time of overall survival T and loss of follow-up C given the treatment assignment $W = w$.

5. Dependent observations

5.1. Survival models

Up to this point, the models we have discussed have been built upon the assumption of independent observations. However, this assumption may not hold in situations where observations are clustered, such as for multi-center trials where observations from the same center are generally correlated.

Mixed-effects proportional hazards model

In order to address this challenge, we can leverage a flexible mixed-effects proportional hazards model as proposed by Tamási *et al.* (2022). This approach extends the previously discussed smooth transformation models by conditioning on an unobserved cluster-specific random effect $R = r$ that accounts for the dependence within clusters,

$$S_w(t \mid R = r) \quad = \quad \mathbb{P}(T > t \mid W = w, R = r) = \exp\left\{-\exp\left[h(t) + \beta w + r\right]\right\}.$$

This formulation provides a fully parametric version of the Cox proportional hazards model (2), incorporating multivariate normally distributed random effects with a zero mean and variance τ^2. The treatment effect β is interpreted as a log-hazard ratio conditional on unobserved random effects. For more in-depth information on likelihood-based inference, see Tamási and Hothorn (2021) and Tamási *et al.* (2022).

Marginalised proportional hazards model

Furthermore, we can explore the model proposed by Barbanti and Hothorn (2024), where the marginal survivor functions are characterised by models (2), while the correlations within clusters are captured by a joint multivariate normal distribution. This joint modeling approach yields a marginalised formulation for the random intercept model,

$$S_w(t) \quad = \quad \mathbb{P}(T > t \mid W = w) = \exp\left\{ - \exp\left[\frac{h(t) + \beta w}{\sqrt{\xi^2 + 1}} \right] \right\}.$$

Here, ξ^2 represents the variance of a cluster-specific intercept. Using this framework, it becomes possible to quantify the treatment effect using the marginal hazard ratio given by $\exp\left(\beta/\sqrt{\xi^2+1}\right)$.

Further details on the models, including likelihood-based inference, can be found in Barbanti and Hothorn (2024).

5.2. Application

To estimate mixed-effects smooth transformation models (Section 5.1.1) we can use the **tramME** package (Tamási 2024; Tamási and Hothorn 2021), available from CRAN:

```
R> install.packages("tramME")
R> library("tramME")
```

Including a random-intercept for the block used in the randomisation, which is a combination of the centers $j = 1, \ldots 88$ and the stratum s ($j \times s =$ `Block`) in the model, we can account for potential correlation between patients from the same block. The corresponding mixed-effects proportional hazards model can be estimated using the `CoxphME()` function.

```
R> CoxphME(iDFS ~ randarm + (1 | Block), data = CAOsurv)
```

Coefficient	Estimate	Std. Error	95%-Wald CI
β	-0.235	0.107	-0.445 to -0.026
τ^2	0.071		
Log-Likelihood			
-2'241.99			

The model provides an estimate of the log-hazard ratio β, which is conditional on the unobserved random effects. The estimated hazard ratio of $\exp(\hat{\beta}) = 0.790$ and corresponding 95%-confidence intervals indicate prolonged disease-free survival time in the experimental arm. The estimated variance of the random effect R is relatively small, with $\hat{\tau}^2 = 0.071$. We can further examine the corresponding marginal estimates of the survivor curves or related measures by integrating out the random effects (for more details, see Tamási and Hothorn (2021)). The corresponding marginal survivor curves for patients from all blocks are depicted in Figure 7.

The estimated mixed-effects proportional hazards model using `CoxphME()` translates the analysis conducted in Rödel *et al.* (2015) into the smooth transformation model framework. The

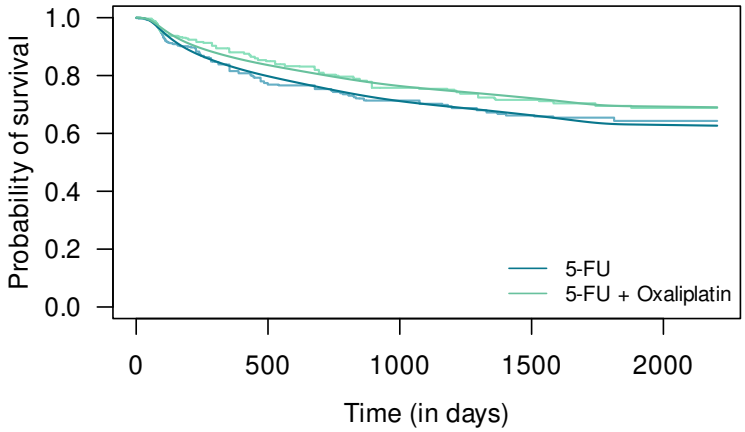


Figure 7: Mixed-effects proportional hazards model. Marginal survivor curves obtained from integrating out the random effects from the model, along with the non-parametric Turnbull estimates.

aim of the primary analysis of Rödel *et al.* (2015) was to fit a Cox model for clustered observations estimating the treatment effect and corresponding standard errors. However, at the time of the primary analysis it was not feasible to estimate the mixed-effects Cox model while accounting for interval-censoring. Fortunately, here the discrepancies between the reported results from the model ignoring interval-censoring and the fitted one, accounting for it, are practically negligible.

To obtain a marginal hazard ratio we can estimate the model that captures the dependence within clusters using a joint multivariate normal distribution (Section 5.1.2), which can be fitted using the `mtram()` function from the **tram** package. Estimation is straightforward for completely exact or interval-censored outcomes within a cluster. Since `iDFS` comprises a mix of different types of outcomes (within each cluster, event times can be either all exact or all censored, see Formulae 2.5 and 2.6 of Barbanti and Hothorn (2024)), we handle exact event times by treating them as interval-censored, accomplished by adding a 4-day window (stored in the object `iDFS2` of class '`Surv`', see Section 5 of the `mtram` package vignette (Barbanti and Hothorn 2025) for details). Again, for the package vignette only we stop the optimisation early

```
R> mtram(Coxph(iDFS2 ~ randarm, data = CAOsurv),
+    formula = ~ (1 | Block), data = CAOsurv, optim = fastoptH)
```

Coefficient	Estimate	Std. Error	95%-Wald CI
β	−0.235	0.107	−0.445 to −0.024
ξ	0.182	0.117	−0.047 to 0.411

Log-Likelihood
−2'047.82

The corresponding estimate of the marginal hazard ratio is $\exp(\hat{\beta}/\sqrt{\hat{\xi}^2+1}) = 0.794$ with empir-

ical 95%-confidence intervals from 0.647 to 0.979. The results indicate that the hazards for patients in the experimental arm is reduced by 0.794 compared to the hazards in the control arm, regardless of the block.

6. Personalised medicine

In the context of personalised medicine, our focus now turns towards modeling heterogeneous treatment effects to capture a more individualised response to treatment. By fitting models with log-hazard ratios that vary with age, we move beyond a global treatment effect, to assess differences in treatment response across age groups.

6.1. Survival models

Age-varying hazards model

To detect varying treatment effects based on age we can employ models which estimate an age-varying hazard ratio $\exp(\beta(\text{age}))$ (Tamási 2025),

$$S_w(t) \quad = \quad \mathbb{P}(T > t \mid W = w, \text{Age} = \text{age}) = \exp\left\{-\exp\left[h(t) + \beta(\text{age})w\right]\right\}.$$

This formulation aligns with the model estimated in the analysis of Hofheinz *et al.* (2018).

Such models could be further extended to additionally capture variations in baseline risks across age by including an age-dependent log-cumulative baseline hazard function: $\log(\Lambda_0(t \mid \text{age})) = h(t \mid \text{age}) = \boldsymbol{a}(t)^\top \boldsymbol{\vartheta} + \beta_0(\text{age})$.

Tree-based age-varying hazards model

Furthermore, for estimating heterogeneous treatment effects, tree-based Cox models can also be employed (Korepanova *et al.* 2019; Seibold *et al.* 2018),

$$S_w(t \mid \text{Age} = \text{age}) \quad = \quad \exp\left\{-\exp\left[h(t \mid \text{age}) + \beta(\text{age})w\right]\right\},$$

allowing to partition both the log-cumulative baseline hazard $\log(\Lambda_0(t \mid \text{age})) = h(t \mid \text{age}) = \boldsymbol{a}(t)^\top \boldsymbol{\vartheta}(\text{age})$ and the treatment effect $\beta(\text{age})$ with respect to different age groups. In contrast to the model in Section 6.1.1, here both the log-cumulative baseline hazard and the log-hazard ratio β depend on age, in this case via an age-structured tree.

6.2. Application

The hazards model featuring an age-varying treatment effect (Section 6.1.1) can be fitted using the **tramME** package (Tamási 2024).

```
R> CoxphME(iDFS ~ randarm + s(age, by = as.ordered(randarm),
+      fx = TRUE, k = 6), data = CAOsurv)
```

The model estimates a global treatment effect and an additional smooth effect for age in the experimental arm, specified as an unpenalized term (using `fx = TRUE`) to match the approach

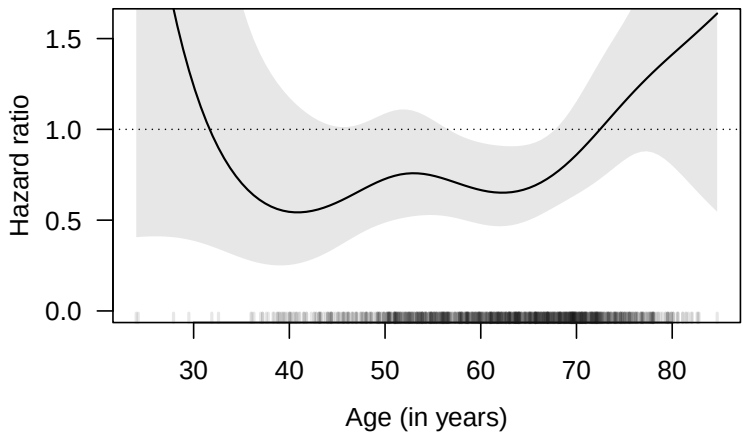


Figure 8: Age-varying hazards model. Hazard ratio and corresponding 95%-confidence interval estimated by the model shown along age. The log-likelihood of the corresponding model is −2'237.63.

used in Hofheinz *et al.* (2018). From the model terms, one can compute an age-varying hazard ratio $\exp(\beta(\text{age}))$.

The estimated age-varying hazard ratio curve, shown in Figure 8, indicates that the experimental treatment is more effective than the control treatment for patients aged $40 - 70$ years, while for older patients the control treatment reduces the hazard compared to the experimental treatment. The corresponding 95%-confidence interval, however, is notably wide and mostly overlaps with a hazard ratio of 1.

Fitting a model partitioning the log-cumulative baseline hazards and treatment effect by age, a survival tree (Section 6.1.2) can be estimated using the **trtf** package (Hothorn 2025c; Hothorn and Zeileis 2021).

```
R> install.packages("trtf")
R> library("trtf")
```

```
R> trafotree(Coxph(iDFS ~ randarm, data = CAOsurv),
+    formula = iDFS ~ randarm | age, data = CAOsurv,
+    control = ctree_control(teststat = "maximum", alpha = .1,
+      minbucket = 40))
```

The survivor functions corresponding to the terminal nodes of the estimated tree are shown in Figure 9. The results again indicate that the experimental treatment is more effective for younger patients, while the control treatment is more effective for older patients. This result is also in line with the one previously obtained by Hofheinz *et al.* (2018).

7. Other extensions

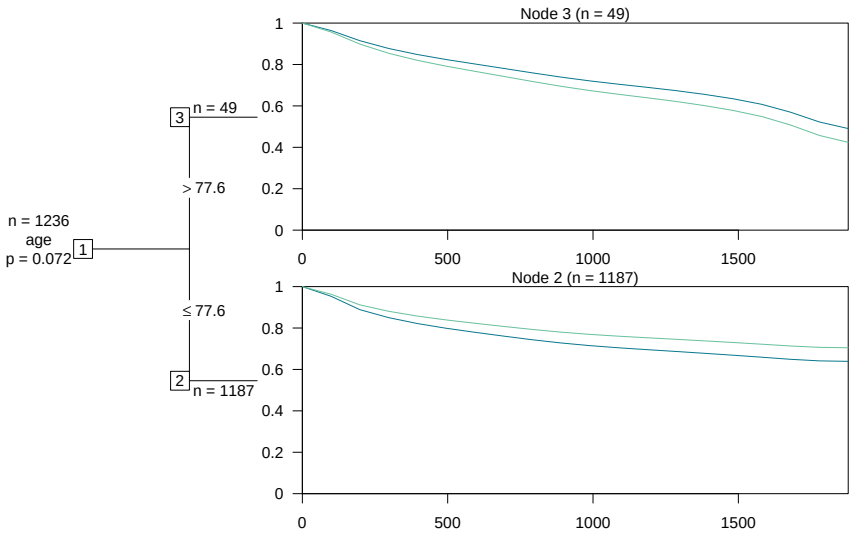


Figure 9: Tree-based age-varying hazards model. Survival tree depicting the estimated survivor curves of the age-groups corresponding to the terminal nodes of the partitioned hazards model. The corresponding in-sample log-likelihood is −2'236.21.

7.1. Frailty proportional hazards model

In cases where the assumption of a homogeneous study population falls short, frailty models offer a valuable alternative. These models account for unobserved heterogeneity in scenarios where the study population comprises individuals with varying baseline risks (Balan and Putter 2020).

To handle such scenarios in the framework of smooth transformation models, the approach discussed by McLain and Ghosh (2013) can be employed. The corresponding frailty proportional hazards model introduces an unobservable multiplicative frailty effect A on the hazard, with corresponding conditional survivor function

$$S_w(t \mid A = a) \quad = \quad \mathbb{P}(T > t \mid W = w, A = a) = \exp\left\{-\exp\left[h(t) + \log(a) + \beta w\right]\right\}.$$

The model implies that the proportional hazards assumption, $\Lambda_1(t|a)/\Lambda_0(t|a) = \lambda_1(t|a)/\lambda_0(t|a) = \beta$, holds conditional on frailty $A = a$. The frailty A specifies a latent random term, assumed to have a certain non-negative distribution, such as the gamma, inverse Gaussian or positive stable distribution (Hougaard 1986), which, for identifiability, is scaled such that $\mathbb{E}(A) = 1$. The proportional hazards model with gamma frailty can be fitted in **tram**, using the `Coxph()` function specifying the frailty distribution with `frailty = "Gamma"`.

7.2. Non-parametric likelihood

In this tutorial, we have primarily focused on the implementation of smooth parametrisation for the log-cumulative baseline hazard function h. Nevertheless, it is important to highlight that researchers also have the option to utilize the discussed models that incorporate a non-parametric version of the transformation function h in package **tram**, should they wish to do so. The corresponding non-parametric transformation function h is specified in terms

of $\boldsymbol{a}(t_k)^\top \boldsymbol{\vartheta} = \vartheta_k$, where a parameter ϑ_k is assigned to each distinct event time t_k with $k = 1, \ldots, K-1$. A head-to-head comparison of the smooth parametrisation and the non-parametric version can be found in Tian *et al.* (2020).

7.3. Link function

Undoubtedly, the proportional hazards model stands as a cornerstone in survival analysis, prominently emerging from specifying the complementary log-log link (the cumulative distribution function of the standard minimum extreme value distribution), wherein h parametrises the log-cumulative baseline hazard function. Nevertheless, it is worth noting that researchers have the option to explore other link functions for all the models shown above, such as the logit link (as also discussed in detail in Royston and Parmar (2002)), or the probit or log-log link.

For example, specifying a flexible proportional odds model "only" requires to change the link function from complementary log-log to logit; such a model can be estimated via

```
R> Colr(iDFS ~ randarm, data = CAOsurv)
```

Coefficient	Estimate	Std. Error	95%-Score CI
β	-0.291	0.124	-0.534 to -0.047
Log-Likelihood	**Likelihood Ratio Test**	**Score Test**	**Permutation Score Test**
$-2'242.06$	$p = 0.019$	$p = 0.019$	$p = 0.019$

This inherent versatility of link functions facilitates to construct alternative models, including mean or odds models, by specifying a probit or logit link respectively. These models are well known from the class of accelerated failure time models (with log-linear h), but extend seamlessly to the more flexible framework of smooth transformation models. Moreover, by selecting the log-log link, it is possible to define a reverse time hazards model. For a comprehensive overview, see Table 1 of Hothorn *et al.* (2018).

7.4. Covariate adjustment

While the models above focus on estimating the treatment effect, they can naturally extend to incorporate further covariates $\boldsymbol{x}$. For example, in the time-varying hazards model (Section 3.1.5), age can be incorporated into the linear predictor as follows:

```
R> Coxph(iDFS | randarm ~ age, data = CAOsurv)
```

Additionally, penalised covariate effects can be estimated by maximizing the L_1- or L_2-penalised log-likelihood using the **tramnet** package (Kook 2023; Kook and Hothorn 2021).

Moreover, conditional transformation models (Hothorn *et al.* 2014), which accommodate complex, non-linear covariate effects, can be estimated using package **tbm** (Hothorn 2025b, 2020b).

7.5. Sample size estimation and simulation

The framework of smooth transformation models can also be valuable for researchers involved in designing new studies. Simulating from the illustrated models (using `simulate()`) offers a

straightforward approach for tasks such as sample size estimation. Because the transformation function h is smooth, it is relatively simple to invert this function numerically, such that it becomes possible to apply probability integral transforms for generating new event times from a fitted model analogously to Bender *et al.* (2005). As an example, we might want to generate data for a future trial where 5-FU overall survival is improved by some innovative therapy. We start with fitting a Weibull model to overall survival, conditional on treatment w and age.

```
R> m <- as.mlt(Survreg(OS ~ randarm + age, data = CAOsurv,
+    dist = "weibull", support = c(.1, 80 * 365)))
```

We simulate new survival times T from this conditional distribution for study participants with normally distributed age in a balanced trial, with the covariate values stored in a data frame called `nd`.

A useful feature in **tram** is the ability to change model coefficients on the fly. Here, we change the log-hazard ratio to 0.25 and simulate from this altered Weibull model:

```
R> cf <- coef(m)
R> cf["randarm5-FU + Oxaliplatin"] <- .25
R> coef(m) <- cf
R> nd$T <- as.Surv(simulate(m, newdata = nd, K = 1000))
```

In addition, we simulate censoring times $C \perp\!\!\!\perp T \mid W = w, \text{Age} = \text{age}$ such that 80% of observations will be right-censored (with probabilistic index $\mathbb{P}(T > C \mid W = w, \text{Age} = \text{age}) = 0.8 = \text{logit}^{-1}(1.386)$ (Sewak and Hothorn 2023))

```
R> cf["(Intercept)"] <- cf["(Intercept)"] + qlogis(.8)
R> coef(m) <- cf
R> nd$C <- as.Surv(simulate(m, newdata = nd, K = 1000))
```

and finally compute the potentially right-censored event times for model re-fitting:

```
R> nd$nOS <- with(nd, Surv(time = pmin(T[, "time"], C[, "time"]),
+    event = T[,"time"] < C[,"time"]))
```

8. Discussion

Motivated by the complexities researchers face when navigating various software implementations for survival analysis, we outline the potential of leveraging smooth transformation models (Hothorn *et al.* 2018) in R. Together with related add-on packages such as **tramME** (Tamási 2024) and **trtf** (Hothorn 2025c), the **tram** package provides a unified maximum likelihood framework that seamlessly extends classical survival models to accommodate more complex scenarios, offering a versatile toolkit for survival analysis.

Throughout this tutorial, we present practical strategies for addressing prominent challenges in survival analysis in R, including interval-censored outcomes, non-proportional or crossing hazards, dependent censoring, clustered observations, and heterogeneous treatment effects.

The comparative overview of implementations in Supplementary Material A serves as a validation tool, allowing researchers to compare across multiple packages.

The frameworks' modular architecture further allows individual model components to be combined — for example, covariate-dependent hazard ratios can be paired with random effects using **tramME**. The framework also extends to multiple event times per subject, allowing for the estimation of multivariate survival models via the `Mmlt()` function (Klein *et al.* 2022). This flexibility provides users with a unified toolkit to seamlessly address a wide spectrum of complex survival analysis tasks in R.

The implemented framework also provides the foundation for interesting extensions. For example, the model in Section 3.1.4 could be adapted to handle cured patients, as a fully parametric version of the semi-parametric cure mixture model proposed by Xie *et al.* (2022). In the context of covariate adjustment, extending the implementation to collapsible models similar to Crowther *et al.* (2023) could be explored. Additionally, alternative strategies such as marginalising the hazard ratio, as suggested by Daniel *et al.* (2021), could also be explored further.

Acknowledgements

This work was supported by the Swiss National Science Foundation [grant number 200021_219384].

References

Balan TA, Putter H (2020). "A Tutorial on Frailty Models." *Statistical Methods in Medical Research*, **29**(11), 3424–3454. doi:10.1177/0962280220921889.

Barbanti L, Hothorn T (2024). "A Transformation Perspective on Marginal and Conditional Models." *Biostatistics*, **25**(2), 402–428. doi:10.1093/biostatistics/kxac048.

Barbanti L, Hothorn T (2025). *Some Applications of Marginally Interpretable Linear Transformation Models for Clustered Observations*. R package vignette, version 1.3-0, URL https://doi.org/10.32614/CRAN.package.tram.

Bender R, Augustin T, Blettner M (2005). "Generating Survival Times to Simulate Cox Proportional Hazards Models." *Statistics in Medicine*, **24**(11), 1713–1723. doi:10.1002/sim.2059.

Boruvka A, Cook RJ (2014). ***coxinterval**: Cox-Type Models for Interval-Censored Data*. R package version 1.0, URL https://CRAN.R-project.org/src/contrib/Archive/coxinterval/.

Burke K, MacKenzie G (2017). "Multi-Parameter Regression Survival Modeling: An Alternative to Proportional Hazards." *Biometrics*, **73**(2), 678–686. doi:10.1111/biom.12625.

Cox DR (1972). "Regression Models and Life-Tables." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **34**(2), 187–220. doi:10.1111/j.2517-6161.1972.tb00899.x.

Cox DR (1975). "Partial Likelihood." *Biometrika*, **62**(2), 269–276. doi:10.2307/2335362.

Crowther MJ, Royston P, Clements M (2023). "A Flexible Parametric Accelerated Failure Time Model and the Extension to Time-Dependent Acceleration Factors." *Biostatistics*, **24**(3), 811–831. doi:10.1093/biostatistics/kxac009.

Czado C, Van Keilegom I (2023). "Dependent Censoring Based on Parametric Copulas." *Biometrika*, **110**(3), 721–738. doi:10.1093/biomet/asac067.

Daniel R, Zhang J, Farewell D (2021). "Making Apples From Oranges: Comparing Non-collapsible Effect Estimators And Their Standard Errors After Adjustment for Different Covariate Sets." *Biometrical Journal*, **63**(3), 528–557. doi:10.1002/bimj.201900297.

Deresa NW, Van Keilegom I (2024). "Copula Based Cox Proportional Hazards Models for Dependent Censoring." *Journal of the American Statistical Association*, **119**(546), 1044–1054. doi:10.1080/01621459.2022.2161387.

Farouki RT (2012). "The Bernstein Polynomial Basis: A Centennial Retrospective." *Computer Aided Geometric Design*, **29**(6), 379–419. doi:10.1016/j.cagd.2012.03.001.

Hofheinz RD, Arnold D, Fokas E, Kaufmann M, Hothorn T, Folprecht G, Fietkau R, Hohenberger W, Ghadimi M, Liersch T, Grabenbauer G, Sauer R, Rödel C, Graeven U (2018). "Impact of Age on the Efficacy of Oxaliplatin in the Preoperative Chemoradiotherapy and Adjuvant Chemotherapy of Rectal Cancer: A Post Hoc Analysis of the CAO/ARO/AIO-04 Phase III Trial." *Annals of Oncology*, **29**(8), 1793–1799. doi:10.1093/annonc/mdy205.

Hothorn T (2020a). "Most Likely Transformations: The **mlt** Package." *Journal of Statistical Software*, **92**(1), 1–68. doi:10.18637/jss.v092.i01.

Hothorn T (2020b). "Transformation Boosting Machines." *Statistics and Computing*, **30**, 141–152. doi:10.1007/s11222-019-09870-4.

Hothorn T (2025a). ***mlt**: Most Likely Transformations*. R package version 1.6-7, URL https://CRAN.R-project.org/package=mlt.

Hothorn T (2025b). ***tbm**: Transformation Boosting Machines*. doi:10.32614/CRAN.package.tbm. R package version 0.3-7, URL https://CRAN.R-project.org/package=tbm.

Hothorn T (2025c). ***trtf**: Transformation Trees and Forests*. doi:10.32614/CRAN.package.trtf. R package version 0.4-3, URL https://CRAN.R-project.org/package=trtf.

Hothorn T, Kneib T, Bühlmann P (2014). "Conditional Transformation Models." *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **76**(1), 3–27. doi:10.1111/rssb.12017.

Hothorn T, Möst L, Bühlmann P (2018). "Most Likely Transformations." *Scandinavian Journal of Statistics*, **45**(1), 110–134. doi:10.1111/sjos.12291.

Hothorn T, Siegfried S, Kook L (2025). ***tram**: Transformation Models*. R package version 1.2-4, URL https://CRAN.R-project.org/package=tram.

Hothorn T, Zeileis A (2021). "Predictive Distribution Modeling Using Transformation Forests." *Journal of Computational and Graphical Statistics*, **30**(4), 1181–1196. doi:10.1080/10618600.2021.1872581.

Hougaard P (1986). "Survival Models for Heterogeneous Populations Derived from Stable Distributions." *Biometrika*, **73**(2), 387–396. doi:10.2307/2336215.

Kalbfleisch JD, Prentice RL (2002). *The Statistical Analysis of Failure Time Data*, volume 360. John Wiley & Sons. doi:10.1002/9781118032985.

Kaplan EL, Meier P (1958). "Nonparametric Estimation from Incomplete Observations." *Journal of the American Statistical Association*, **53**(282), 457–481.

Klein N, Hothorn T, Barbanti L, Kneib T (2022). "Multivariate Conditional Transformation Models." *Scandinavian Journal of Statistics*, **49**, 116–142. doi:10.1111/sjos.12501.

Kook L (2023). ***tramnet**: Penalized Transformation Models*. doi:10.32614/CRAN.package.tramnet. R package version 0.0-8, URL https://CRAN.R-project.org/package=tramnet.

Kook L, Hothorn T (2021). "Regularized Transformation Models: The **tramnet** Package." *The R Journal*, **13**(1), 581–594. doi:10.32614/RJ-2021-054.

Korepanova N, Seibold H, Steffen V, Hothorn T (2019). "Survival Forests under Test: Impact of the Proportional Hazards Assumption on Prognostic and Predictive Forests for ALS Survival." *Statistical Methods in Medical Research*. doi:10.1177/0962280219862586.

McLain AC, Ghosh SK (2013). "Efficient Sieve Maximum Likelihood Estimation of Time-Transformation Models." *Journal of Statistical Theory and Practice*, **7**(2), 285–303. doi:10.1080/15598608.2013.772835.

Peng D, MacKenzie G, Burke K (2020). "A Multiparameter Regression Model for Interval-Censored Survival Data." *Statistics in Medicine*, **39**(14), 1903–1918. doi:10.1002/sim.8508.

Rödel C, Graeven U, Fietkau R, Hohenberger W, Hothorn T, Arnold D, Hofheinz RD, Ghadimi M, Wolff HA, Lang-Welzenbach M, Raab HR, Wittekind C, Ströbel P, Staib L, Wilhelm M, Grabenbauer GG, Hoffmanns H, Lindemann F, Schlenska-Lange A, Folprecht G, Sauer R, Torsten Liersch on behalf of the German Rectal Cancer Study Group (2015). "Oxaliplatin Added to Fluorouracil-based Preoperative Chemoradiotherapy and Postoperative Chemotherapy of Locally Advanced Rectal Cancer (the German CAO/ARO/AIO-04 study): Final Results of the Multicentre, Open-label, Randomised, Phase 3 Trial." *The Lancet Oncology*, **16**(8), 979–989. doi:10.1016/S1470-2045(15)00159-X.

Royston P, Parmar MKB (2002). "Flexible Parametric Proportional-hazards and Proportional-odds Models for Censored Survival Data, with Application to Prognostic Modelling and Estimation of Treatment Effects." *Statistics in Medicine*, **21**(15), 2175–2197. doi:10.1002/sim.1203.

Seibold H, Zeileis A, Hothorn T (2018). "Individual Treatment Effect Prediction for Amyotrophic Lateral Sclerosis Patients." *Statistical Methods in Medical Research*, **27**(10), 3104–3125. doi:10.1177/0962280217693034.

Sewak A, Hothorn T (2023). "Estimating Transformations for Evaluating Diagnostic Tests with Covariate Adjustment." *Statistical Methods in Medical Research*, **32**(7), 1403–1419. doi:10.1177/09622802231176030.

Siegfried S, Kook L, Hothorn T (2023). "Distribution-Free Location-Scale Regression." *The American Statistician*, **77**(4), 345–356. doi:10.1080/00031305.2023.2203177.

Tamási B (2024). ***tramME**: Transformation Models with Mixed Effects*. doi:10.32614/CRAN.package.tramME. R package version 1.0.7, URL https://CRAN.R-project.org/package=tramME.

Tamási B (2025). "Mixed-Effects Additive Transformation Models with the R Package **tramME**." *Journal of Statistical Software*, **114**(11). doi:10.18637/jss.v114.i11.

Tamási B, Crowther M, Puhan MA, Steyerberg E, Hothorn T (2022). "Individual Participant Data Meta-Analysis with Mixed-Effects Transformation Models." *Biostatistics*, **23**(4), 1083–1098. doi:10.1093/biostatistics/kxab045.

Tamási B, Hothorn T (2021). "**tramME**: Mixed-Effects Transformation Models Using Template Model Builder." *The R Journal*, **13**(1), 581–594. doi:10.32614/RJ-2021-075.

Therneau TM (2024). ***survival**: Survival Analysis*. doi:10.32614/CRAN.package.survival. R package version 3.8-3, URL https://CRAN.R-project.org/package=survival.

Therneau TM, Grambsch PM (2000). *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, New York, U.S.A.

Tian Y, Hothorn T, Li C, Harrell Jr FE, Shepherd BE (2020). "An Empirical Comparison of Two Novel Transformation Models." *Statistics in Medicine*, **39**(5), 562–576. doi:10.1002/sim.8425.

Turnbull BW (1976). "The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data." *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **38**(3), 290–295.

Wei LJ (1992). "The Accelerated Failure Time Model: A Useful Alternative to the Cox Regression Model in Survival Analysis." *Statistics in Medicine*, **11**(14-15), 1871–1879. doi:10.1002/sim.4780111409.

Xie C, Huang X, Li R, Pisters PWT (2022). "A Flexible-Hazards Cure Model With Application to Patients With Soft Tissue Sarcoma." *Statistics in Medicine*, **41**(29), 5698–5714. doi:10.1002/sim.9588.

A. Comparative overview of R implementations

In the following, we compare the implementations of the models from the **tram** (Hothorn *et al.* 2025) and **tramME** (Tamási 2024) packages shown in the tutorial, with alternative models available in various established R packages from CRAN.

We fit the models to the primary endpoint of disease-free survival, which comprises a mixture of exact times and right- and interval-censored event times (encoded in `iDFS`, an object of class '`Surv`'). In order to further compare the models with other implementations that cannot handle interval-censored outcomes, we treat the interval-censored observations as if they were observed exactly (encoded in `DFS`, an object of class '`Surv`').

We contrast treatment effect estimates (Estimate) and corresponding standard errors (Std. Error) estimated by the fitted models. It is important to note the difference in interpretation of the estimates (Interpretation). Additionally, we provide the in-sample log-likelihood (logLik) of the fitted models, with penalised or semi-parametric/partial likelihoods highlighted in grey.

A.1. Weibull models

The **survival** package (Therneau 2024b), the **icenReg** package (Anderson-Bergman 2024, 2017) and the **flexsurv** package (Jackson 2024) provide alternative implementations of Weibull models. Both the **survival** and **icenReg** package (specifying `model = "aft"`) implement accelerated failure time Weibull models, where the effect can be interpreted as log-acceleration factor (log-AF). An alternative parametrisation of such models is in terms of proportional hazards Weibull models, estimating log-hazard ratios (log-HRs), instead. This is how the model, fitted by `Survreg()`, is implemented in the **tram** package (Hothorn *et al.* 2025), which can be directly compared to the analogous parametrisation in the **icenReg** (specifying `model = "ph"`) and the **flexsurv** package (with `dist = "weibullPH"`). All Weibull models can handle interval-censoring, owing to the parametric nature of the models.

The Weibull models can be fitted to the *interval-censored outcomes* as follows:

```
R> tram::Survreg(iDFS ~ randarm, data = CAOsurv, dist = "weibull")
R> icenReg::ic_par(iDFS ~ randarm, data = CAOsurv, dist = "weibull",
+    model = "ph")
R> flexsurv::flexsurvreg(iDFS ~ randarm, data = CAOsurv,
+    dist = "weibullPH")
R> survival::survreg(iDFS ~ randarm, data = CAOsurv, dist = "weibull")
R> icenReg::ic_par(iDFS ~ randarm, data = CAOsurv, dist = "weibull",
+    model = "aft")
```

Function	Package	Interpretation	Estimate	Std. Error	logLik
`Survreg`	**tram**	log-HR	−0.229	0.106	−2'281.17
`ic_par`	**icenReg**	log-HR	−0.229	0.106	−2'281.17
`flexsurvreg`	**flexsurv**	log-HR	−0.229	0.106	−2'281.17
`survreg`	**survival**	log-AF	0.312	0.146	−2'281.17
`ic_par`	**icenReg**	log-AF	0.312	0.146	−2'281.17

As expected, all packages provide equivalent model fits.

A.2. Flexible proportional hazards models

Flexible versions of the proportional hazards model are implemented in several packages, of which the following accommodate interval-censored outcomes. The **rstpm2** package (Clements *et al.* 2025; Liu *et al.* 2016) and the **flexsurv** package provide parametric versions of the model by using splines (analogously to the approach discussed by Royston and Parmar (2002)). We set `k = 3` for the number of knots in the spline for `flexsurvspline()` from the **flexsurv** package Alternatively, the **icenReg** package (Anderson-Bergman 2024) provides a semi-parametric implementation of the model that can handle interval-censoring.

The corresponding models can be fitted to the *interval-censored outcomes* as follows:

```
R> tram::Coxph(iDFS ~ randarm, data = CAOsurv)
R> rstpm2::stpm2(Surv(time = iDFStime, time2 = iDFStime2,
+    event = iDFSevent, type = "interval") ~ randarm, data = CAOsurv)
R> flexsurv::flexsurvspline(iDFS ~ randarm, data = CAOsurv, k = 3)
R> icenReg::ic_sp(iDFS ~ randarm, data = CAOsurv, model = "ph")
```

Function	Package	Interpretation	Estimate	Std. Error	logLik
`Coxph`	**tram**	log-HR	−0.231	0.107	−2'242.25
`stpm2`	**rstpm2**	log-HR	−0.232	0.107	−2'250.48
`flexsurvspline`	**flexsurv**	log-HR	−0.231	0.106	−2'254.34
`ic_sp`	**icenReg**	log-HR	−0.230	–	−1'977.29

The models fit similarly across all four packages. Due to the fact that the computations of the standard errors of `ic_sp()` from the **icenReg** package rely on computationally expensive bootstrap sampling, we did not report any standard errors for this approach. Also the log-likelihood (in grey) of the semi-parametric model from the **icenReg** is not comparable to the otherwise fully parametric implementations.

The **ICsurv** package (McMahan and Wang 2022) could also potentially handle interval-censored event times. The **TransModel** package (Zhou *et al.* 2022b,a), featuring an alternative implementation of linear transformation model, could also serve as an interesting comparator. However, we encountered difficulties and errors when trying to fit the model using these two packages.

In scenarios where *interval-censoring* is not taken into account, there are several other implementations available for fitting corresponding models. The `coxph()` function from the **survival** package provides a semi-parametric approach for exact or right-censored observations (Therneau 2024b). (Note, that again the likelihood of the fitted model is not comparable to the other fully parametric models and thus marked in grey). The **rms** package (Harrell Jr 2025) implements a semi-parametric model, in line with the model from package **survival**.

```
R> tram::Coxph(DFS ~ randarm, data = CAOsurv)
R> survival::coxph(DFS ~ randarm, data = CAOsurv)
R> rms::cph(DFS ~ randarm, data = CAOsurv)
```

Function	Package	Interpretation	Estimate	Std. Error	logLik
`Coxph`	**tram**	log-HR	−0.230	0.106	−3'264.89
`coxph`	**survival**	log-HR	−0.228	0.106	−2'430.66
`cph`	**rms**	log-HR	−0.228	0.106	−2'430.66

A.3. Stratified proportional hazards models

For comparing stratified flexible proportional hazards models we can again utilize the model from the **rstpm2**, which employ stratified spline-basis functions. The model can be fitted to the *interval-censored event times* as follows

```
R> tram::Coxph(iDFS | strat ~ randarm, data = CAOsurv)
R> rstpm2::stpm2(Surv(time = iDFStime, time2 = iDFStime2,
+      event = iDFSevent, type = "interval") ~ randarm +
+      strata(strat), data = CAOsurv)
```

Function	Package	Interpretation	Estimate	Std. Error	logLik
`Coxph`	**tram**	log-HR	−0.228	0.107	−2'213.94
`stpm2`	**rstpm2**	log-HR	−0.220	0.107	−2'242.88

The results from the two models are practically similar.

Now, *ignoring interval-censoring*, we can, once again, contrast the implementation of the semi-parametric models from the **survival** package and the **rms** package:

```
R> tram::Coxph(DFS | strat ~ randarm, data = CAOsurv)
R> survival::coxph(DFS ~ randarm + strata(strat), data = CAOsurv)
R> rms::cph(DFS ~ randarm + strat(strat), data = CAOsurv)
```

Function	Package	Interpretation	Estimate	Std. Error	logLik
`Coxph`	**tram**	log-HR	−0.228	0.107	−3'234.58
`coxph`	**survival**	log-HR	−0.222	0.107	−2'089.54
`cph`	**rms**	log-HR	−0.222	0.107	−2'089.54

The three model fits are practically equivalent.

We can proceed to compare the stratified version of the Weibull model, for which we also will ignore interval-censoring due to the fact that the utilised **eha** package (Broström 2024) lacks support for interval-censored data. Additionally, it is worth highlighting that there is a distinction from the model fitted using `survreg()` from the **survival** package (Therneau 2024b). This model only stratifies the scale parameter of the Weibull distribution, whereas the models from the **eha** package and the **tram** package estimate both strata-dependent scale and shape terms. The `survreg` function from the **survival** package fits an accelerated failure time Weibull model, while the **eha** package implements a proportional hazards Weibull model, analogously to the `Survreg()` implementation from the **tram** package. The models can be fitted to the exact event times, *ignoring interval-censoring*, as follows

```
R> tram::Survreg(DFS | strat ~ randarm, data = CAOsurv)
R> eha::phreg(DFS ~ randarm + strata(strat), data = CAOsurv)
R> survival::survreg(DFS ~ randarm + strata(strat), data = CAOsurv)
```

Function	Package	Interpretation	Estimate	Std. Error	logLik
`Survreg`	**tram**	log-HR	−0.219	0.107	−3'277.35
`phreg`	**eha**	log-HR	−0.219	0.107	−3'277.35
`survreg`	**survival**	log-AF	0.274	0.133	−3'280.87

The fit of the `survreg` model from **survival** package is slightly different. In contrast, the parametrisation and fits of the model from the **eha** and the **tram** are equivalent.

A.4. Non-proportional hazards models

To the best of our knowledge, there is currently no implementation available that estimates an analogous model to the flexible non-proportional (location-scale) hazards model implemented in **tram**.

However we can contrast the implementation of the less-flexible Weibull model with the **gamlss** package (Stasinopoulos and Rigby 2024, 2007) using the `WEI2` distribution and the **gamlss.cens** package (Stasinopoulos *et al.* 2023) to account for *interval-censoring*.

```
R> tram::Survreg(iDFS ~ randarm | randarm, data = CAOsurv,
+    remove_intercept = FALSE)
R> gamlss::gamlss(formula = iDFS ~ randarm, sigma.fo = ~ randarm,
+    family = gamlss.cens::cens(family = "WEI2", type = "interval"),
+    data = CAOsurv[, c("iDFS", "randarm")],
+    control = gamlss.control(n.cyc = 300, trace = FALSE))
```

Since the scale term in the **tram** package and the **gamlss** package are parameterised differently, we only show the estimates and standard errors of the location parameter below.

Function	Package	Estimate	Std. Error	logLik
`Survreg`	**tram**	−0.847	0.536	−2'280.47
`gamlss`	**gamlss**	−0.948	0.542	−2'280.53

The two implementations provide almost equivalent model fits.

The **mpr** package (Burke 2022) also offers an implementation for a non-proportional Weibull model, it, however, does not support interval-censored data. Thus we fit the models *ignoring interval-censoring*.

```
R> tram::Survreg(DFS ~ randarm | randarm, data = CAOsurv,
+    remove_intercept = FALSE)
R> mpr::mpr(DFS ~ list(~ randarm, ~ randarm), data = CAOsurv)
```

Again, we only show the estimates and standard errors of the location parameter below.

Function	Package	Estimate	Std. Error	logLik
`Survreg`	**tram**	−0.976	0.568	−3'290.43
`mpr`	**mpr**	−0.975	0.567	−3'290.43

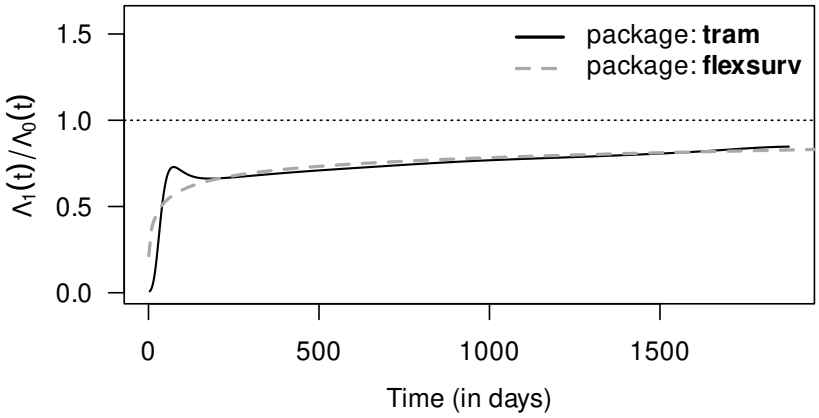
The two implementations also provide equivalent model fits.

A.5. Time-varying hazards model

We can compare the time-varying hazards model from the **tram** and the **flexsurv** package (Jackson 2024) which allows to estimate time-varying treatment effects. We start by examining the models for the *interval-censored event times*.

```
R> tram::Coxph(iDFS | randarm ~ 1, data = CAOsurv)
R> flexsurv::flexsurvspline(iDFS ~ randarm + gamma1(randarm) +
+      gamma2(randarm), data = CAOsurv, k = 3)
```

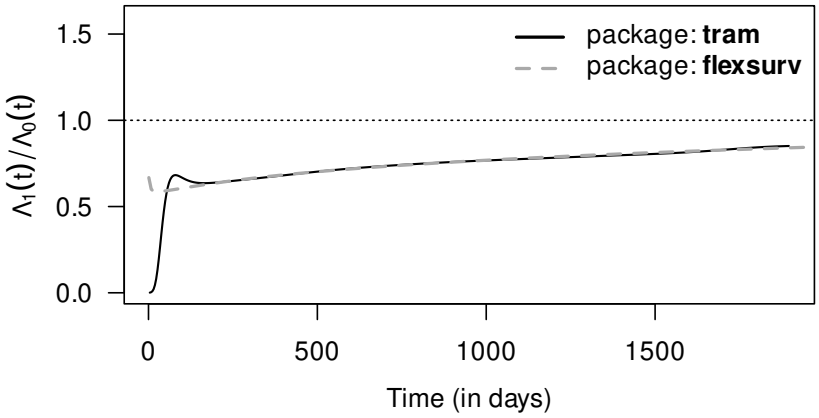
The in-sample log-likelihood is −2'252.95 for the **flexsurv** model and −2'240.21 for the **tram** model. The estimated time-varying ratios of the cumulative hazards are shown in the plot below.



We will now explore the same models *ignoring interval-censoring*.

```
R> tram::Coxph(DFS | randarm ~ 1, data = CAOsurv)
R> flexsurv::flexsurvspline(DFS ~ randarm + gamma1(randarm) +
+      gamma2(randarm), data = CAOsurv, k = 3)
```

The in-sample log-likelihood is −3'267.27 for the **flexsurv** model and −3'262.54 for the **tram** model, with the computed time-varying ratios of the cumulative hazards shown below.



The time-varying effects estimated from `DFS` show good agreement, the ratios slightly differ when the models are estimated on the interval-censored data (`iDFS`).

A.6. Mixed-effects proportional hazards models

The implementation of a mixed-effects proportional hazards model with flexible log-cumulative

baseline hazards for interval-censored event times is unique in the **tramME** package (Tamási 2024; Tamási and Hothorn 2021). While the **rstpm2** package also accommodates interval-censored event times, we were not able to fit the corresponding mixed-effects model to our data.

Thus, to contrast the models with other implementations we, again, need to *ignore interval-censoring*. We can then compare the fitted model with the fully parametric spline-based version from the **rstpm2** package (Clements *et al.* 2025; Liu *et al.* 2017) and the semi-parametric model estimated by the **coxme** package (Therneau 2024a), employing Gaussian random effects using a Laplace approximation (Ripatti and Palmgren 2000).

```
R> tramME::CoxphME(DFS ~ randarm + (1 | Block), data = CAOsurv)
R> rstpm2::stpm2(Surv(DFStime, DFSevent) ~ randarm, data = CAOsurv,
+    cluster = "Block", RandDist = "LogN")
R> coxme::coxme(DFS ~ randarm + (1 | Block), data = CAOsurv)
```

Function	Package	Interpretation	Estimate	Std. Error	logLik
`CoxphME`	**tramME**	log-HR	−0.234	0.107	−3'264.66
`stpm2`	**rstpm2**	log-HR	−0.234	0.107	−3'272.86
`coxme`	**coxme**	log-HR	−0.231	0.107	−2'414.48

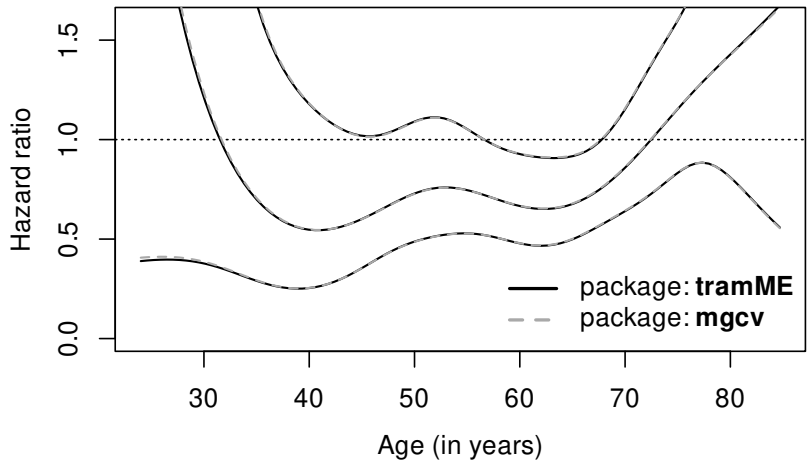
The fitted models from the three packages agree very well.

A.7. Age-varying hazards model

We can compare the age-varying hazards model from package **tramME** (Tamási 2025) to the implementation in the **mgcv** package (Wood 2025; Wood *et al.* 2016) which estimates a smooth Cox model via partial likelihood optimisation. As the model from the **mgcv** package only accommodates right-censored observations we again fit the models *ignoring interval-censoring*.

```
R> tramME::CoxphME(DFS ~ randarm + s(age, by = as.ordered(randarm),
+      fx = TRUE, k = 6), data = CAOsurv)
R> mgcv::gam(DFStime ~ randarm + s(age, by = as.ordered(randarm),
+      fx = TRUE, k = 6), data = CAOsurv, family = cox.ph(),
+    weights = DFSevent)
```

The in-sample log-likelihood of the model from the package **mgcv** is −2'426.04 (partial log-likelihood) and −3'260.25 for the **tramME** model. The estimated age-varying hazard ratios and corresponding 95%-confidence intervals are shown in the plot below.



The hazard ratio curves and confidence intervals estimated by the two packages are practically equivalent.

A.8. Frailty proportional hazards models

For models featuring a gamma frailty, we can contrast implementations using a semi-parametric approach or the spline-based approach from the **rstpm2** package (Liu *et al.* 2017). The `coxph()` model from the **survival** package uses a semi-parametric approach and estimates the frailty term using penalised regression (Therneau *et al.* 2003). The **frailtyEM** (Balan and Putter 2019b,a) and the **frailtypack** package (Rondeau *et al.* 2025, 2012) also feature models with semi-parametric baseline hazards. Again we fit the models *ignoring interval-censoring*.

```
R> tram::Coxph(DFS ~ randarm, data = CAOsurv, frailty = "Gamma")
R> rstpm2::stpm2(Surv(DFStime, DFSevent) ~ randarm, data = CAOsurv,
+    cluster = "id", RandDist = "Gamma")
R> survival::coxph(DFS ~ randarm + frailty(id, distribution = "gamma"),
+    data = CAOsurv)
R> frailtyEM::emfrail(DFS ~ randarm + cluster(id), data = CAOsurv)
R> frailtypack::frailtyPenal(DFS ~ randarm + cluster(id),
+    data = CAOsurv, RandDist = "Gamma", n.knots = 10, kappa = 1)
```

Function	Package	Interpretation	Estimate	Std. Error	logLik
`Coxph`	**tram**	log-HR	−0.230	0.107	−3'264.89
`stpm2`	**rstpm2**	log-HR	−0.685	0.670	−3'264.88
`coxph`	**survival**	log-HR	−0.406	0.159	−1'944.22
`emfrail`	**frailtyEM**	log-HR	−0.384	0.153	−2'430.45
`frailtyPenal`	**frailtypack**	log-HR	−0.660	0.248	−3'259.82

The fitted models vary considerably across packages.

A.9. Flexible proportional odds models

We can compare the fit of the flexible proportional odds model with the implementation in the **rstpm2** package (Clements *et al.* 2025) and package **flexsurv** (Jackson 2024). The `Gprop.odds`

function from package **timereg** (Scheike and Martinussen 2025; Scheike and Zhang 2011) can also estimate a flexible proportional odds model using the partial likelihood, thus we again compare the models *ignoring interval-censoring*.

```
R> tram::Colr(DFS ~ randarm, data = CAOsurv)
R> rstpm2::stpm2(Surv(DFStime, DFSevent) ~ randarm, data = CAOsurv,
+    link.type = "PO")
R> flexsurv::flexsurvspline(iDFS ~ randarm, data = CAOsurv, k = 3,
+    scale = "odds")
R> timereg::Gprop.odds(DFS ~ prop(randarm), data = CAOsurv)
```

Function	Package	Interpretation	Estimate	Std. Error	logLik
`Colr`	**tram**	log-OR	−0.292	0.125	−3'265.48
`stpm2`	**rstpm2**	log-OR	−0.294	0.125	−3'272.44
`flexsurvspline`	**flexsurv**	log-OR	−0.294	0.124	−2'247.78
`Gprop.odds`	**timereg**	log-OR	−0.268	0.125	

The fitted models are practically equivalent among the four packages. Note, that we were not able to retrieve the in-sample log-likelihood from the model object of the **timereg** package and thus do not report it here.

Computational details

- R version 4.5.2 (2025-10-31), `x86_64-pc-linux-gnu`
- Running under: `Ubuntu 24.04.4 LTS`
- Matrix products: default
- BLAS: `/usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3`
- LAPACK: `/usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblasp-r0.3.26.so` ; LAPACK version3.12.0
- Base packages: `base, datasets, graphics, grDevices, grid, methods, parallel, splines, stats, utils`
- Other packages: `ATR 0.1-1, basefun 1.2-5, bdsmatrix 1.3-7, boot 1.3-32, coda 0.19-4.1, coin 1.4-3, colorspace 2.1-2, coxme 2.2-22, doBy 4.7.1, eha 2.11.5, fastGHQuad 1.0.1, flexsurv 2.3.2, frailtyEM 1.0.1, frailtypack 3.8.0, gamlss 5.5-0, gamlss.cens 5.0-7, gamlss.data 6.0-7, gamlss.dist 6.1-1, Hmisc 5.2-4, icenReg 2.0.16, ICsurv 1.0.1, knitr 1.51, libcoin 1.0-10, MASS 7.3-65, mgcv 1.9-4, mlt 1.7-3, mpr 1.0.6, multcomp 1.4-29, mvtnorm 1.3-3, nlme 3.1-168, optimx 2025-4.9, parfm 2.7.8, partykit 1.2-25, Rcpp 1.1.1, rms 8.1-0, rstpm2 1.7.1, SparseGrid 0.8.2, survC1 1.0-3, survival 3.8-6, TH.data 1.1-5, timereg 2.0.7, tram 1.3-2, tramME 1.0.8, TransModel 2.3, trtf 0.4-3, variables 1.1-2, xtable 1.8-4`

- Loaded via a namespace (and not attached): `alabama 2025.1.0`, `assertthat 0.2.1`, `backports 1.5.0`, `base64enc 0.1-6`, `BB 2019.10-1`, `bbmle 1.0.25.1`, `broom 1.0.12`, `checkmate 2.3.4`, `cli 3.6.5`, `cluster 2.1.8.2`, `codetools 0.2-20`, `compiler 4.5.2`, `coneproj 1.23`, `cowplot 1.2.0`, `data.table 1.17.8`, `Deriv 4.2.0`, `deSolve 1.41`, `digest 0.6.39`, `doParallel 1.0.17`, `dplyr 1.1.4`, `evaluate 1.0.5`, `expint 0.2-1`, `expm 1.0-0`, `farver 2.1.2`, `fastmap 1.2.0`, `foreach 1.5.2`, `forecast 9.0.1`, `foreign 0.8-90`, `Formula 1.2-5`, `fracdiff 1.5-3`, `future 1.67.0`, `future.apply 1.20.0`, `generics 0.1.4`, `ggplot2 4.0.2`, `globals 0.19.0`, `glue 1.8.0`, `gridExtra 2.3`, `gtable 0.3.6`, `htmlTable 2.4.3`, `htmltools 0.5.8.1`, `htmlwidgets 1.6.4`, `inum 1.0-5`, `iterators 1.0.14`, `lattice 0.22-9`, `lava 1.8.2`, `lifecycle 1.0.5`, `listenv 0.10.0`, `magrittr 2.0.4`, `marqLevAlg 2.0.8`, `Matrix 1.7-4`, `matrixcalc 1.0-6`, `MatrixModels 0.5-4`, `matrixStats 1.5.0`, `microbenchmark 1.5.0`, `mnormt 2.1.2`, `modelr 0.1.11`, `modeltools 0.2-24`, `msm 1.8.2`, `mstate 0.3.3`, `muhaz 1.2.6.4`, `nloptr 2.2.1`, `nnet 7.3-20`, `numDeriv 2016.8-1.1`, `orthopolynom 1.0-6.1`, `otel 0.2.0`, `parallelly 1.46.1`, `pillar 1.11.1`, `pkgconfig 2.0.3`, `polspline 1.1.25`, `polynom 1.4-1`, `pracma 2.4.6`, `purrr 1.2.1`, `quadprog 1.5-8`, `quantreg 6.1`, `R6 2.6.1`, `rbibutils 2.4.1`, `RColorBrewer 1.1-3`, `Rdpack 2.6.6`, `reformulas 0.4.4`, `rlang 1.1.7`, `rmarkdown 2.30`, `rootSolve 1.8.2.4`, `rpart 4.1.24`, `rstudioapi 0.17.1`, `S7 0.2.1`, `sandwich 3.1-1`, `scales 1.4.0`, `sn 2.1.2`, `SparseM 1.84-2`, `statmod 1.5.1`, `stats4 4.5.2`, `stringi 1.8.7`, `stringr 1.6.0`, `tibble 3.3.1`, `tidyr 1.3.2`, `tidyselect 1.2.1`, `timeDate 4052.112`, `TMB 1.9.19`, `tools 4.5.2`, `urca 1.3-4`, `vctrs 0.7.1`, `withr 3.0.2`, `xfun 0.56`, `zoo 1.8-15`

References

Anderson-Bergman C (2017). "**icenReg**: Regression Models for Interval Censored Data in R." *Journal of Statistical Software*, **81**(12), 1–23. doi:10.18637/jss.v081.i12.

Anderson-Bergman C (2024). ***icenReg**: Regression Models for Interval Censored Data*. doi:10.32614/CRAN.package.icenReg. R package version 2.0.16, URL https://CRAN.R-project.org/package=icenReg.

Balan TA, Putter H (2019a). "**frailtyEM**: An R Package for Estimating Semiparametric Shared Frailty Models." *Journal of Statistical Software*, **90**(7). doi:10.18637/jss.v090.i07.

Balan TA, Putter H (2019b). ***frailtyEM**: Fitting Frailty Models with the EM Algorithm*. doi:10.32614/CRAN.package.frailtyEM. R package version 1.0.1, URL https://CRAN.R-project.org/package=frailtyEM.

Broström G (2024). ***eha**: Event History Analysis*. doi:10.32614/CRAN.package.eha. R package version 2.11.5, URL https://CRAN.R-project.org/package=eha.

Burke K (2022). ***mpr**: Multi-Parameter Regression (MPR)*. doi:10.32614/CRAN.package.mpr. R package version 1.0.6, URL https://CRAN.R-project.org/package=mpr.

Clements M, Liu XR, Christoffersen B (2025). ***rstpm2**: Smooth Survival Models, Including Generalized Survival Models*. doi:10.32614/CRAN.package.rstpm2. R package version 1.6.9, URL https://CRAN.R-project.org/package=rstpm2.

Harrell Jr FE (2025). ***rms**: Regression Modeling Strategies*. doi:10.32614/CRAN.package.rms. R package version 8.0-0, URL https://CRAN.R-project.org/package=rms.

Hothorn T, Siegfried S, Kook L (2025). ***tram**: Transformation Models*. R package version 1.2-4, URL https://CRAN.R-project.org/package=tram.

Jackson C (2024). ***flexsurv**: Flexible Parametric Survival and Multi-State Models*. doi:10.32614/CRAN.package.flexsurv. R package version 2.3.2, URL https://CRAN.R-project.org/package=flexsurv.

Liu XR, Pawitan Y, Clements M (2016). "Parametric and Penalized Generalized Survival Models." *Statistical Methods in Medical Research*, **27**(5), 1531–1546. doi:10.1177/0962280216664760.

Liu XR, Pawitan Y, Clements M (2017). "Generalized Survival Models for Correlated Time-To-Event Data." *Statistics in Medicine*, **36**(29), 4743–4762. doi:10.1002/sim.7451.

McMahan CS, Wang L (2022). ***ICsurv**: Semiparametric Regression Analysis of Interval-Censored Data*. doi:10.32614/CRAN.package.ICsurv. R package version 1.0.1, URL https://CRAN.R-project.org/package=ICsurv.

Ripatti S, Palmgren J (2000). "Estimation of Multivariate Frailty Models Using Penalized Partial Likelihood." *Biometrics*, **56**(4), 1016–1022. doi:10.1111/j.0006-341X.2000.01016.x.

Rondeau V, Gonzalez JR, Mazroui Y, Mauguen A, Diakite A, Laurent A, Lopez M, Krol A, Sofeu CL, Dumerc J, Rustand D, Chauvet J, Le Coent Q, Pierlot R, Etzkorn L, Dinart D, Orué A (2025). ***frailtypack**: Shared, Joint (Generalized) Frailty Models; Surrogate Endpoints*. doi:10.32614/CRAN.package.frailtypack. R package version 3.7.0, URL https://CRAN.R-project.org/package=frailtypack.

Rondeau V, Marzroui Y, Gonzalez JR (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. doi:10.18637/jss.v047.i04.

Royston P, Parmar MKB (2002). "Flexible Parametric Proportional-hazards and Proportional-odds Models for Censored Survival Data, with Application to Prognostic Modelling and Estimation of Treatment Effects." *Statistics in Medicine*, **21**(15), 2175–2197. doi:10.1002/sim.1203.

Scheike T, Martinussen T (2025). ***timereg**: Flexible Regression Models for Survival Data*. doi:10.32614/CRAN.package.timereg. R package version 2.0.7, URL https://CRAN.R-project.org/package=timereg.

Scheike TH, Zhang MJ (2011). "Analyzing Competing Risk Data Using the R **timereg** Package." *Journal of Statistical Software*, **38**(2), 1–15. doi:10.18637/jss.v038.i02.

Stasinopoulos DM, Rigby RA (2007). "Generalized Additive Models for Location Scale and Shape (GAMLSS) in R." *Journal of Statistical Software*, **23**(7), 1–46. doi:10.18637/jss.v023.i07.

Stasinopoulos M, Rigby B, Mortan N, Seipp A (2023). ***gamlss.cens**: Fitting an Interval Response Variable Using **gamlss.family** Distributions*. doi:10.32614/CRAN.package.gamlss.cens. R package version 5.0-7, URL https://CRAN.R-project.org/package=gamlss.cens.

Stasinopoulos M, Rigby R (2024). ***gamlss**: Generalized Additive Models for Location Scale and Shape*. doi:10.32614/CRAN.package.gamlss. R package version 5.4-22, URL https://CRAN.R-project.org/package=gamlss.

Tamási B (2024). ***tramME**: Transformation Models with Mixed Effects*. doi:10.32614/CRAN.package.tramME. R package version 1.0.7, URL https://CRAN.R-project.org/package=tramME.

Tamási B (2025). "Mixed-Effects Additive Transformation Models with the R Package **tramME**." *Journal of Statistical Software*, **114**(11). doi:10.18637/jss.v114.i11.

Tamási B, Hothorn T (2021). "**tramME**: Mixed-Effects Transformation Models Using Template Model Builder." *The R Journal*, **13**(1), 581–594. doi:10.32614/RJ-2021-075.

Therneau TM (2024a). ***coxme**: Mixed Effects Cox Models*. doi:10.32614/CRAN.package.coxme. R package version 2.2-22, URL https://CRAN.R-project.org/package=coxme.

Therneau TM (2024b). ***survival**: Survival Analysis*. doi:10.32614/CRAN.package.survival. R package version 3.8-3, URL https://CRAN.R-project.org/package=survival.

Therneau TM, Grambsch PM, Pankratz VS (2003). "Penalized Survival Models and Frailty." *Journal of Computational and Graphical Statistics*, **12**(1), 156–175. URL https://www.jstor.org/stable/1391074.

Wood S (2025). ***mgcv**: Mixed GAM Computation Vehicle with Automatic Smoothness Estimation*. doi:10.32614/CRAN.package.mgcv. R package version 1.9-3, URL https://CRAN.R-project.org/package=mgcv.

Wood S, Pya N, Säfken B (2016). "Smoothing Parameter and Model Selection for General Smooth Models (with Discussion)." *Journal of the American Statistical Association*, **111**, 1548–1575. doi:10.1080/01621459.2016.1180986.

Zhou J, Zhang J, Lu W (2022a). "**TransModel**: An R Package for Linear Transformation Model with Censored Data." *Journal of Statistical Software*, **101**(9), 1–12. doi:10.18637/jss.v101.i09.

Zhou J, Zhang J, Lu W (2022b). ***TransModel**: Fit Linear Transformation Models for Right Censored Data*. doi:10.32614/CRAN.package.TransModel. R package version 2.3, URL https://CRAN.R-project.org/package=TransModel.

Affiliation:

Sandra Siegfried, Bálint Tamási, and Torsten Hothorn
Institut für Epidemiologie, Biostatistik und Prävention
Universität Zürich
Hirschengraben 84, CH-8001 Zürich, Switzerland
`Siegfried.Sandra@protonmail.ch`, `Torsten.Hothorn@uzh.ch`