On Estimating Survival and Hazard Ratios Using One Time-Scale When Cohort Data Have Multiple Time-Scales: A Simulation Study

Nurgul Batyrbekova (✉ nurgul.batyrbekova@ki.se)

Karolinska Institute

Hannah Bower

Karolinska Institute

Paul Dickman

Karolinska Institute

Robert Szulkin

SDS Life Science AB

Paul C. Lambert

University of Leicester

Therese M-L. Andersson

Karolinska Institute

Research Article

Keywords: multiple time-scales, flexible parametric models, simulation study, bias, survival, hazard ratios

Posted Date: January 26th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-151021/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

RESEARCH

On estimating survival and hazard ratios using one time-scale when cohort data have multiple time-scales: a simulation study

Nurgul Batyrbekova^{1,2*}, Hannah Bower³, Paul Dickman¹, Robert Szulkin², Paul C. Lambert^{1,4} and Therese M-L. Andersson^{1,10}

*Correspondence:

nurgul.batyrbekova@ki.se

¹Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Full list of author information is available at the end of the article

Abstract

Background: When estimating survival functions and hazard ratios during the analysis of cohort data, we often choose one time-scale, such as time-on-study, as the primary time-scale, and include a fixed covariate, such as age at entry, in the model. However, we rarely consider the possibility of simultaneous effects of multiple time-scales on the hazard function.

Methods: In a simulation study, within the framework of flexible parametric models, we investigate whether relying on one time-scale and fixed covariate as proxy for the second time-scale is sufficient in capturing the true survival functions and hazard ratios when there are actually two underlying time-scales.

Result: We demonstrate that the one-time-scale survival models appeared to approximate well the survival proportions, however, large bias was observed in the log hazard ratios if the covariate of interest had interactions with the second time-scale or with both time-scales.

Conclusion: We recommend to exercise caution and encourage fitting models with multiple time-scales if it is suspected that the cohort data have underlying non-proportional hazards on the second time-scale or both time-scales.

Keywords: multiple time-scales; flexible parametric models; simulation study; bias; survival; hazard ratios

Introduction

Survival analysis of time-to-event data can include several time-scales, such as time since diagnosis, treatment duration, attained age or calendar time. Multiple time-

scales can be seen in situations, where the event rate may vary along more than one time-scale. For example, breast cancer incidence depends on attained age as well as time since first childbirth; in many chronic diseases such as diabetes, the risk of complications is contingent on time since diagnosis and attained age; the risk of getting an infection in intensive care unit may depend on length of stay and calendar time [1, 2].

When analysing cohort data, we may be aware of the presence of two or more time-scales that may simultaneously affect the rates of an outcome. Time-on-study and attained age are the most frequently encountered time-scales. However, the joint effects of both time-scales are rarely considered. Instead, one time-scale, such as time-on-study, is often chosen as the main time-scale for the baseline hazard, while the effect of the second time-scale is not modelled, rather a time-fixed covariate, such as age at entry is included in the model.

The usual approach to jointly modelling two time-scales is to treat the main time-scale as continuous, and split the data into very short intervals along the second time-scale, thus assuming constant effects within each interval. The obtained data can then be analysed with a Cox model, or a parametric model. The flexible parametric models have become a popular alternative in analysing survival data, as they can capture complex hazard shapes, and enable smooth predictions for relative and absolute effects [3, 4, 5]. A Poisson regression model can also be used if the observations are further split along the first time-scale in addition to time-intervals on the second time-scale. This may lead to very fine splitting of the observed data to uphold the assumption of the constant hazard rate within each interval.

In large population-based register studies, splitting observations into very short intervals along two or more time-scales can become unmanageable with current technology. Time-splitting can further exacerbate the challenges of fitting very complex models. The absence of a well documented standard software to jointly model multiple time-scales can be an additional deterrent to fitting multiple time-scales models. These limitations play a role in our preference to model how the rate changes as a function of only one time-scale. Unfortunately, we rarely consider the potential bias in estimates from a misspecified model that ignores simultaneous effects of more than one time-scale.

¹ In this study, within the framework of flexible parametric models, we aim to¹
²investigate whether modelling one time-scale while adjusting for a constant proxy²
³for the second time-scale suffices in different situations when the hazard rates in³
⁴the data depend on two time-scales. In situations where these models appear to⁴
⁵be insufficient, we also aim to assess bias in survival and hazard ratio estimates.⁵
⁶Assessment of choosing an optimal time-scale is out of scope of this study as there⁶
⁷has been a considerable attention paid to the choice of a time-scale when analysing⁷
⁸cohort data [6, 7, 8, 9], as well as combining time-scales into a unique optimal single⁸
⁹time-scale [10, 11, 12].⁹

¹¹Methods¹¹

¹²Models with one or more time-scales¹²

¹³At least one time origin is inherent in the analysis of time-to-event data. Examples¹³
¹⁴include the start of study, onset of disease or date of birth. It is common to model¹⁴
¹⁵the log hazard function as a function of the time-scale of interest when analysing¹⁵
¹⁶time-to-event data. The frequency of an event of interest from the time origin can¹⁶
¹⁷dictate which time-scale(s) to use when modelling the baseline hazard function. For¹⁷
¹⁸example, in cancer patients, the rate of deaths increases from onset of the disease,¹⁸
¹⁹making the time from onset, or rather its proxy, the time from diagnosis, as the¹⁹
²⁰time-scale of interest. However, attained age, which is the time from birth, also plays²⁰
²¹an important role in the rates of death. We cannot know for certain whether one²¹
²²time-scale should be preferred over the other in the analysis of time-to-event data as²²
²³there are many other risk factors that can influence the time at risk. Conventionally,²³
²⁴one of the two time-scales is chosen as the primary time-scale in the baseline hazard²⁴
²⁵function without further consideration that modelling both time-scales may help in²⁵
²⁶capturing the true hazard function.²⁶

²⁷ A general model for the log hazard function with only one time-scale, t_1 , such as²⁷
²⁸time from diagnosis, can be expressed as:²⁸

$$\begin{aligned} \log(h(t_1; \gamma, \beta)) &= \log(h_0(t_1; \gamma)) + \mathbf{x}\beta \\ &= g(t_1; \gamma) + \mathbf{x}\beta, \end{aligned} \quad (1) \quad \begin{matrix} 29 & 30 & 31 \end{matrix}$$

³² where h_0 is the baseline hazard function, γ is the corresponding parameter vector³²
³³ and $\mathbf{x}$ is a vector of covariates with associated log hazard ratios β . In the Cox model,³³

¹the baseline hazard is not modelled directly due to the use of the partial likelihood¹
²[13], whereas in the Poisson regression it is common to model the behaviour of²
³the baseline hazard by categorising t_1 through function g into time-intervals, say³
⁴monthly or yearly intervals. If a continuous format of the time-scale is preferred then⁴
⁵ g can be any polynomial function or a spline function, which allows for a smooth⁵
⁶non-linear function. 6

⁷ In situations where it is known that the hazard rate may depend on two time-⁷
⁸scales, such as time from diagnosis and attained age, the usual approach is to include⁸
⁹the time from diagnosis, t_1 , in the baseline hazard and use age at entry, a_0 , as a⁹
¹⁰constant proxy for the second time-scale, as shown in the following model: 10

$$\begin{aligned} \log(h(t_1; \gamma, \beta^*)) &= \log(h_0(t_1; \gamma)) + \mathbf{x}^* \beta^* \\ &= g(t_1; \gamma) + f(a_0; \beta_{a_0}) + \mathbf{x} \beta. \end{aligned} \quad (2) \quad \begin{matrix} 11 \\ 12 \\ 13 \end{matrix}$$

¹⁴In this model a_0 is part of the covariates $\mathbf{x}^*$, and the relative effect of a_0 , modelled¹⁴
¹⁵by function f , is part of the vector of the log hazard ratios β^* . Similar to g , function¹⁵
¹⁶ f is usually specified as a categorical function (such as five-yearly intervals), or as¹⁶
¹⁷a smoothing spline function. To allow for time-dependent effects of age at diagnosis¹⁷
¹⁸and of any x_p ($p = 1, \dots, P$) covariate in $\mathbf{x}$, model (2) can include interaction terms¹⁸
¹⁹with t_1 : 19

$$\begin{aligned} \log(h(t_1; \gamma, \beta_{a_0}, \gamma_{a_0}, \beta, \gamma_{x_p})) &= g(t_1; \gamma) + f(a_0; \beta_{a_0}) + r(g_{int}(t_1) f_{int}(a_0); \gamma_{a_0}) \\ &\quad + \mathbf{x} \beta + \sum_{p=1}^P g_{x_p}(t_1; \gamma_{x_p}) x_p. \end{aligned} \quad (3) \quad \begin{matrix} 20 \\ 21 \\ 22 \\ 23 \end{matrix}$$

²⁴where P is the number of covariates, excluding age, with time-dependent effects. 24

²⁵Now, a model that includes both time from diagnosis, t_1 , and attained age, t_2 , in
²⁶the baseline hazard function, h_0 , assuming proportional hazards, can be written as:
²⁷27

$$\begin{aligned} \log(h(t_1, t_2; \gamma, \psi, \beta)) &= \log(h_0(t_1, t_2; \gamma, \psi)) + \mathbf{x} \beta \\ &= g(t_1; \gamma) + f(t_2; \psi) + \mathbf{x} \beta. \end{aligned} \quad (4) \quad \begin{matrix} 28 \\ 29 \end{matrix}$$

³⁰Furthermore, model (4) can also incorporate interaction between two time-scales t_1
³¹and t_2 : 31

$$\log(h(t_1, t_2; \gamma, \psi, \tau, \beta)) = g(t_1; \gamma) + f(t_2; \psi) + r(g_{int}(t_1) f_{int}(t_2); \tau) + \mathbf{x} \beta. \quad (5) \quad \begin{matrix} 32 \\ 33 \end{matrix}$$

¹If necessary, this model can be extended to include more time-scales, and interaction¹
²between time-scales and other covariates.²

³ Depending on the origin of the time-scales, often one time-scale can be expressed³
⁴in terms of the other. This may not always be true if the time-scale is defined in⁴
⁵different units to time units. For example, failure rates in vehicles may be primarily⁵
⁶measured in terms of the mileage rather than age, whereas for aircraft the time-⁶
⁷scales can include the number of flights, the time in air and calendar time. However,⁷
⁸in the medical setting, such as our example, using the time of diagnosis as the time⁸
⁹origin, and a_0 as an offset term, we can specify attained age, t_2 , in terms of t_1 and⁹
¹⁰ a_0 , therefore transforming model (4) into¹⁰

$$\log(h(t_1, t_2; \gamma, \psi, \beta)) = g(t_1; \gamma) + f(t_1 + a_0; \psi) + \mathbf{x}\beta. \quad (6)^{12}$$

¹⁴Non-proportional hazards as well as interactions between time-scales can be accom-¹⁴
¹⁵modated by including interactions between the covariates and the time-scales:¹⁵

$$\begin{aligned} \log(h(t_1, t_2; \gamma, \psi, \tau, \beta, \gamma_{x_p}, \psi_{x_q})) = & g(t_1; \gamma) + f(t_1 + a_0; \psi) + \\ & + r(g_{int}(t_1)f_{int}(t_1 + a_0); \tau) + \\ & + \mathbf{x}\beta + \sum_{p=1}^P g_{x_p}(t_1; \gamma_{x_p})x_p + \\ & + \sum_{q=1}^Q f_{x_q}(t_1 + a_0; \psi_{x_q})x_q. \end{aligned} \quad (7)^{18}$$

²² In this study, within the framework of flexible parametric models, our focus is on²²
²³errors in predictions when fitting models as shown in equations (2), (3) with one²³
²⁴time-scale while adjusting for a second time-scale as a constant covariate, when the²⁴
²⁵true underlying rates depend on multiple time-scales as shown in equations (6), (7).²⁵

Flexible parametric models

²⁷ Flexible parametric survival models (FPMs) are parametric models that use re-²⁷
²⁸stricted cubic splines to model the baseline hazard function as well as time-²⁸
²⁹dependent effects on the log hazard scale or log cumulative hazard scale [3, 4, 5]. A²⁹
³⁰flexible parametric model on the log hazard scale with one time-scale that includes³⁰
³¹interactions with time, is written as:³¹

$$\log(h(t; \gamma_0, \mathbf{k}_0, \beta, \gamma_p, \mathbf{k}_p)) = s(\log(t); \gamma_0, \mathbf{k}_0) + \mathbf{x}\beta + \sum_{p=1}^P x_p s(\log(t); \gamma_p, \mathbf{k}_p), \quad (8)^{33}$$

¹where h is a hazard function over time-scale t , $\mathbf{x}$ is a vector of covariates with asso-¹
²ciated log hazard ratios β , with additional p^{th} time-dependent effect for covariate²
³ x_p , with $p = 1, \dots, P$. The baseline hazard is represented by the restricted cubic³
⁴spline function, $s(\log(t); \gamma_0, \mathbf{k}_0)$, with knot location vector, $\mathbf{k}_0$, and associated coef-⁴
⁵ficient vector γ_0 from the spline expansion [5]. Users need to determine the number⁵
⁶of knots for the spline expansion of $\log(t)$ and for each of the spline expansions for⁶
⁷the time-dependent effects. Interactions between the covariates and the restricted⁷
⁸cubic spline function of $\log(t)$ is omitted to fit a model with proportional hazards.⁸
⁹It should be noted that it is common to model $\log(t)$ in FPMs on the log hazard⁹
¹⁰scale, but other functional forms of time t can be used as well. ¹⁰

¹¹ ¹¹ ¹²Simulation strategy ¹²

¹³This simulation study was conducted to evaluate bias in estimates of survival func-¹³
¹⁴tions and hazard ratios when modelling hazard rates using FPMs with one time-¹⁴
¹⁵scale given that the true hazard rates depend on two time-scales. For this purpose,¹⁵
¹⁶we simulated cohort data loosely based on haematological cancer patients who re-¹⁶
¹⁷lapsed after transplantation, and were followed from relapse until death. We si-¹⁷
¹⁸multaneously simulated two time-scales – time since relapse and time since initial¹⁸
¹⁹transplantation. In the interest of parsimony, we chose to exclude calendar time and¹⁹
²⁰attained age, which should not be ignored in the clinical study of death rates in²⁰
²¹these patients. To the simulated data we fitted FPMs with time since relapse as the²¹
²²time-scale, and used the time since transplantation until relapse as a constant proxy²²
²³for the time since transplantation, while also allowing for possible non-proportional²³
²⁴hazards. ²⁴

²⁵ ²⁵ ²⁶Simulation design and data generation ²⁶

²⁷We simulated 1000 datasets with a sample size of 1000 from each of the seven data-²⁷
²⁸generating mechanisms (DGMs) with and without interaction terms with time-²⁸
²⁹scales presented in Table 1. Each data-generating mechanism simulated time-to-²⁹
³⁰event data based on two time-scales: the first time-scale, t , time (up to 10 years)³⁰
³¹from relapse until death or censoring, and the second time-scale, $t + c_0$, time from³¹
³²transplantation, where c_0 represents the time from transplantation until relapse,³²
³³and was generated from a uniform distribution $U(0, 5)$. Notation c_0 for offset is³³
³⁴used instead of a_0 from equations (2)-(7) to make it clear that these offsets are

¹from different time-scales. We used *Bernoulli*(0.8) for simulation of the single bi-¹
²nary covariate x . The chosen probability distributions may not be realistic within²
³the context of relapsed patients after transplantation, but they provide a useful³
⁴foundation for the purpose of this study. 4

⁵ Different DGMs with and without interactions with time-scales were chosen to 5
⁶reflect the wide spectrum of situations that we usually encounter when analysing 6
⁷cohort data. Scenarios S_1, S_2 generated data with proportional hazards for x ; S_3, S_4 7
⁸simulated interaction between x and t , while S_5, S_6 simulated interaction between 8
⁹ x and $t + c_0$; S_7 generated data with both interaction terms: between x and t , and 9
¹⁰between x and $t + c_0$. Of the seven data-generating mechanisms, three scenarios 10
¹¹(S_2, S_4, S_6) also simulated interactions between the two time-scales. 11

¹² For illustration of the DGMs, Figure 1 displays the survival proportions and haz- 12
¹³ard rates for $x = 0, 1$ as well as the log hazard ratios ($x = 1$ versus $x = 0$) during 13
¹⁴follow-up time t , which is the time from relapse until death or censoring, given re- 14
¹⁵lapses occurred at $c_0 = 0.5, 4$ years after transplantation, respectively. For scenarios 15
¹⁶ S_5, S_6 , it should be noted that the log hazard ratios are not technically indepen- 16
¹⁷dent of the first time-scale t due to the mathematical relationship between the first 17
¹⁸time-scale and the second time-scale ($t_2 - t_1 = (t + c_0) - t = c_0$). This can be seen 18
¹⁹in declining log hazard ratios over follow-up time t . 19

²²*Analyses of simulated data* 22

²³ To each of the simulated datasets, we fitted flexible parametric models on the log 23
²⁴hazard scale with one time-scale, t , while adjusting for a constant c_0 . For the re- 24
²⁵stricted cubic spline expansion of $\log(t)$, we chose six knots (five degrees of free- 25
²⁶dom), and three knots (two degrees of freedom) for the restricted cubic splines of 26
²⁷the time-dependent effects. The choice of the number of knots was not of interest 27
²⁸in our analyses, and according to previous studies, estimates from FPMs are not 28
²⁹sensitive to the number of knots as long as sufficient number of knots is selected 29
³⁰[14, 15, 16]. Furthermore, the knots positions, as is most common, were chosen ac- 30
³¹cording to the equally-spaced centiles of the distribution of event times, with the 31
³²boundary knots at the first and last event times. The effect of c_0 was also modelled 32
³³with restricted cubic splines, with three degrees of freedom. 33

¹ For each two-time-scale DGM we aimed to have a conceptually corresponding¹
² model with one time-scale that might be a sufficient model to capture the effects of²
³ both time-scales. Table 2 provides formulae for the fitted models and indication of³
⁴ correspondence with the DGMs. ⁴

⁵ Survival proportions for x and log hazard ratios were predicted from each fitted⁵
⁶ model over time-scale t for specific values of c_0 . Average values of bias (the mean⁶
⁷ difference between the predicted values and true values) and 95% Monte-Carlo⁷
⁸ confidence intervals (MC CI) were evaluated at $t = 1, 3, 5$ years and $c_0 = 0.5, 4$ ⁸
⁹ years. ⁹

¹⁰ When analysing real-world data, the underlying data generating mechanism is¹⁰
¹¹ unknown, and the AIC [17] and the BIC [18] are often used to guide in model¹¹
¹² selection. In order to mimic this model fitting procedure we also estimated AIC and¹²
¹³ BIC for each fitted model. We calculated the percentage of times the fitted models¹³
¹⁴ showed lowest AIC or BIC relative to other models when fitted to the same dataset¹⁴
¹⁵ from each DGM. We used the number of events in calculation of the BIC [19]. ¹⁵

¹⁶ All the results from the analyses including interactive figures with average survival¹⁶
¹⁷ predictions and average log hazard ratios for all follow-up time t for each fitted¹⁷
¹⁸ model for each simulating scenario can be found in the Supplementary Material¹⁸
¹⁹ <https://nurbatyr.github.io/Suppl-material-fitting-1ts-to-mult-ts/>. ¹⁹

²⁰ For reference, to mirror the true models used for simulating data, we fitted flexible²⁰
²¹ parametric models that jointly modelled both time-scales using three degrees of²¹
²² freedom for $\log(t)$ and for $t + c_0$. Since these models are close approximations of the²²
²³ true models, they are only presented in the Supplementary Material. ²³

²⁵ Results ²⁵

²⁶ Figures 2 - 4 present the bias in predicted survival functions and the log hazard²⁶
²⁷ ratios at follow-up times $t = 1, 3, 5$ (years since relapse) and $c_0 = 0.5, 4$ (time of²⁷
²⁸ relapse after transplantation) for the fitted models for the DGMs. ²⁸

³⁰ Scenarios with proportional hazards ³⁰

³¹ For two simpler scenarios S_1, S_2 , which simulated data with proportional hazards³¹
³² for x , the absolute bias was low (< 0.06) in the survival estimates from all fitted³²

³³ ^[1]Supplementary material may also be found in the online version of this article. ³³

¹models. In particular, very low absolute values (<0.005) were seen in fitted models¹
² $M2$, $M4$ and $M6$, which included interaction between t and c_0 . Larger absolute²
³bias was observed in the log hazard ratios, where only models $M1$ and $M2$, with³
⁴proportional hazards for x , showed close estimates to the true values (absolute bias⁴
⁵ <0.005). These models were aimed to correspond to S_1 and S_2 , respectively. The⁵
⁶AIC and BIC were both seen to also support models $M1$, $M2$ most of the time in⁶
⁷respect to these scenarios, as shown in Table 3. 7

⁸

⁸

⁹*Scenarios with interactions with the first time-scale* 9

¹⁰For scenarios S_3, S_4 , which generated data with interaction between x and the¹⁰
¹¹first time-scale t , the largest bias in survival proportions and the log hazard ra-¹¹
¹²tios was seen in the proportional hazards models. For example, model $M1$ in sce-¹²
¹³nario S_4 showed a bias of -0.08 (MCSE 0.001) in survival proportions for $x = 0$ at¹³
¹⁴ $t = 1$, $c_0 = 0.5$ and 0.39 (MCSE 0.0026) for the log hazard ratio at $t = 5$. In contrast,¹⁴
¹⁵model $M4$, which conceptually corresponds to S_4 , showed the lowest absolute bias in¹⁵
¹⁶survival proportions in both S_3 and S_4 (<0.003 , <0.006 , respectively). These were¹⁶
¹⁷closely followed by estimates from models $M3$ and $M7$. Very low bias (<0.05) was¹⁷
¹⁸also observed in the log hazard ratios from models $M3$, $M4$ and $M7$. Furthermore,¹⁸
¹⁹model $M4$ showed the lowest AIC 92.2% of the time compared to other models¹⁹
²⁰when fitting data generated by S_4 , and model $M3$ had the lowest AIC in over 67%²⁰
²¹of the fitted samples from S_3 . 21

²²

²²

²³*Scenarios with interactions with the second time-scale* 23

²⁴In comparison to other scenarios, larger absolute bias was seen in both survival²⁴
²⁵proportions and the log hazard ratios from all fitted models for scenarios S_5, S_6 ²⁵
²⁶which simulated interactions with the second time-scale $t + c_0$. Models $M5, M6$ ²⁶
²⁷that included interaction $x \cdot c_0$, and conceptually corresponded to scenarios S_5, S_6 ,²⁷
²⁸respectively, showed low bias (<0.09) in survival predictions. However, model $M7$ ²⁸
²⁹that included both interaction $x \cdot \log(t)$ and interaction $x \cdot c_0$, showed the lowest²⁹
³⁰absolute bias (<0.05) in survival predictions. Furthermore, only model $M7$ was³⁰
³¹seen to have the closest estimates to the true log hazard ratios, particularly for³¹
³² $t < 5$. However, bias increased with increasing follow-up time, as shown in Figures³²
³³10.2, 11.2 in the Supplementary Material. In addition, model $M7$ showed the lowest³³

¹AIC in 83.7% of the fitted samples from S_5 , and only in 19.4% of the fitted samples¹
²from S_6 .²

³
⁴*Scenario with interactions with both time-scales*⁴

⁵For scenario S_7 , which included both interaction terms between x and the two time-⁵
⁶scales, model $M7$ was conceptually the closest model. This model had the lowest⁶
⁷absolute bias (<0.02) in survival predictions compared to other models. It also⁷
⁸showed the closest log hazard ratios to the true values within the first five years⁸
⁹of follow-up time, but bias increased with increasing follow-up time beyond five⁹
¹⁰years as shown in Figure 12.2 in the Supplementary Material. The remaining fitted¹⁰
¹¹models had large absolute bias in both survival proportions and the log hazard¹¹
¹²ratios (>0.1 , >1.1 , respectively). The lowest AIC was seen for model $M7$ in 835 of¹²
¹³the fitted 1000 datasets generated by scenario S_7 .¹³

¹⁴**Discussion**¹⁴

¹⁵The purpose of this simulation study was to assess the performance of fitted models¹⁵
¹⁶with one time-scale (time-on-study) while incorporating time-of-entry on the second¹⁶
¹⁷time-scale as a covariate when in reality data depend jointly on both time-scales.¹⁷
¹⁸To this end, we simulated seven scenarios where we specified two time-scales, and¹⁸
¹⁹evaluated bias in seven fitted flexible parametric models for each scenario. Although¹⁹
²⁰we focused on fitting flexible parametric models with one time-scale to the simu-²⁰
²¹lated data, we expect similar results from fitting Cox or Poisson regression models²¹
²²provided the data are split into very short intervals and a smooth function fitted to²²
²³these intervals.²³

²⁴The majority of the fitted models with one time-scale that approximately corre-²⁴
²⁵sponded to the data-generating mechanisms, showed on average very close estimates²⁵
²⁶to the true survival proportions and hazard ratio estimates, and most of the time²⁶
²⁷were seen to have lowest AIC. Models that included an interaction between main²⁷
²⁸time-scale and the constant from the second time-scale ($c_0 \cdot \log(t)$), while control-²⁸
²⁹ling for proportional hazards or non-proportional hazards for x showed most robust²⁹
³⁰results if data had proportional or non-proportional hazards for x on the first time-³⁰
³¹scale, respectively.³¹

³²All models underperformed when estimating the log hazard ratios from data with³²
³³interaction between x and the second time-scale $t + c_0$. It is highly likely that a³³

¹researcher encountering data similar to data from S_5, S_6 , may choose models that¹
²include an interaction term between x and the main time-scale, such as models $M3$,²
³ $M4$ or $M7$, which include such an interaction. However, this can lead to misleading³
⁴interpretation that the variation in the effect of x on the hazard rates is solely due to⁴
⁵the main time-scale (t) rather than the underlying second time-scale ($t + c_0$). Even⁵
⁶though model $M7$ that includes both interaction $x \cdot \log(t)$ and interaction $x \cdot c_0$ was⁶
⁷seen to have very close survival predictions to the true values, the bias in the log⁷
⁸hazard ratios was still visibly large, especially for follow-up time beyond five years.⁸
⁹We can never be safe from such pitfalls in the analysis of real-world data. With our⁹
¹⁰simulation study, we hope to raise awareness of such pitfalls and direct attention¹⁰
¹¹towards more nuanced analysis with cautious approach to modeling time-scales. It¹¹
¹²should be noted that the simulated scenarios were designed to have joint strong¹²
¹³effects from both time-scales, and therefore we do not expect to observe as severe¹³
¹⁴bias in situations with less strong effects. 14

¹⁵ 15
¹⁶As in the majority of simulation studies, we only looked at limited number of¹⁶
¹⁷scenarios, and made simplifying assumptions. In particular, we simulated data using¹⁷
¹⁸only two time-scales with one type of functional form for both time-scales and¹⁸
¹⁹interactions with time-scales. It would be of interest to investigate the performance¹⁹
²⁰of models fitted to the data generated using different time-scales with different²⁰
²¹functional forms as well as different lengths of follow-up time. In this study we²¹
²²focused on the estimation of the effect of x , and did not explore the estimation of²²
²³the effect of c_0 , which is of great interest in clinical practice. Furthermore, we did²³
²⁴not investigate possible situations of smaller population size. 24

²⁵ 25
²⁶The growing complexity of disease progression and increasing length of follow-up²⁶
²⁷time in chronic disease and cancer patients require more advanced survival mod-²⁷
²⁸els that account for outcomes on multiple time-scales and can provide more reliable²⁸
²⁹estimates with respect to different time-points on either of the time-scales. Unfortu-²⁹
³⁰nately, it is rarely possible to know with confidence whether the hazard rates depend³⁰
³¹on more than one time-scale even if the cohort data may contain other variables³¹
³²measured in units of time. If empirical evidence suggests that the hazard rates are³²
³³dependent on multiple time-scales, then users are encouraged to use models with³³
³³multiple time-scales and, if necessary, allow for non-proportional hazards on either 33

¹of the time-scales. Using the AIC or the BIC may be helpful in selecting the most¹
²appropriate model.²

³
⁴Conclusion⁴

⁵Through this study, we aimed to demonstrate the importance of careful consider-⁵
⁶ation when modelling with one time-scale, especially when cohort data with two⁶
⁷time-scales have interactions between covariates and at least one of the time-scales.⁷

⁸ When modelling hazard rates dependent on two time-scales, where one time-scale⁸
⁹can be expressed as a function of the other time-scale and a constant offset, the flex-⁹
¹⁰ible parametric models with one time-scale appear to approximate well the survival¹⁰
¹¹proportions and log hazard ratios, provided the models correctly accommodate the¹¹
¹²proportional or non-proportional hazards on the first time-scale as well as including¹²
¹³interaction between the first time-scale and the constant c_0 (time-of-entry on the¹³
¹⁴second time-scale). On the other hand, if the cohort data have interactions between¹⁴
¹⁵the covariates and the second time-scale or both time-scales then caution should be¹⁵
¹⁶exercised. Using models with one time-scale with non-proportional hazards for the¹⁶
¹⁷main covariate x and interaction $x \cdot c_0$ may provide good approximation of survival¹⁷
¹⁸proportions but it may still lead to bias in the log hazard ratios. Users may find¹⁸
¹⁹the AIC useful when making model choices in practice, but this is not guaranteed¹⁹
²⁰to select the most appropriate model [15].²⁰

²¹
²²Abbreviations²²

²³ AIC: Akaike information criterion; BIC: Bayesian information criterion; DGM: data-generating mechanism; FPM:
²⁴flexible parametric survival models; MC CI: Monte-Carlo confidence interval;²³

²⁴ Acknowledgements²⁴

²⁵NB and RS are employed by SDS Life Science and their time within this project was funded by SDS Life Science,
²⁶however, SDS Life Science had no involvement in the study.²⁵

²⁷ Authors' contributions²⁷

²⁸ TA, HB and NB designed the study. NB performed the data analysis and drafted the manuscript. HB, TA, PCL, PD
²⁹and RS supported the analysis, and provided revisions and improvements to the manuscript. All authors have read
³⁰and approved the final manuscript.²⁸

³¹ Funding³¹

³² This work was funded via the Swedish Cancer Society (grant number: 19 0102 Pj) and the Swedish Research
³³Council (grant numbers: 2019-01965, 2019-00227).³²

³⁴ Availability of data and materials³⁴

³⁵ All the data were simulated for this study. The program files used for simulating data can be obtained from the
³⁶corresponding author upon request.³⁵

¹ Ethics approval and consent to participate	1
² Not applicable.	2
³ Competing interests	3
⁴ The authors declare that they have no competing interests.	4
⁵ Consent for publication	5
⁶ Not applicable.	6
⁷ Author details	7
¹ Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. ² SDS Life	
⁸ Science AB, Stockholm, Sweden. ³ Division of Clinical Epidemiology, Department of Medicine Solna, Karolinska	8
⁹ Institutet, Stockholm, Sweden. ⁴ Biostatistics Research Group, Department of Health Sciences, University of	9
Leicester, Leicester, UK.	
	10
References	
¹¹ 1. Albrektsen, G., Heuch, I., Hansen, S., Kvåle, G.: Breast cancer risk by age at birth, time since birth and time	11
¹² intervals between births: exploring interaction effects. <i>British journal of cancer</i> 92 (1), 167–175 (2005). doi:10.1038/sj.bjc.6602302	12
¹³ 2. Wolkewitz, M., Cooper, B.S., Palomar-Martinez, M., Alvarez-Lerma, F., Olaechea-Astigarraga, P., Barnett,	13
¹⁴ A.G., Schumacher, M.: Multiple time scales in modeling the incidence of infections acquired in intensive care	14
units. <i>BMC medical research methodology</i> 16 (1) (2016). doi:10.1186/s12874-016-0199-y	
¹⁵ 3. Royston, P., Parmar, M.K.B.: Flexible parametric proportional-hazards and proportional-odds models for	15
¹⁶ censored survival data, with application to prognostic modelling and estimation of treatment effects. <i>Statistics</i>	16
¹⁷ in <i>Medicine</i> 21 (15), 2175–2197 (2002). doi:10.1002/sim.1203. https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.1203	17
¹⁸ 4. Royston, P., Lambert, P.C.: Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model. StataCorp LP, ??? (2011). https://EconPapers.repec.org/RePEc:tsj:spbook:fpsaus	18
¹⁹ 5. Crowther, M.J., Lambert, P.C.: A general framework for parametric survival analysis. <i>Statistics in Medicine</i>	19
²⁰ 33 (30), 5280–5297 (2014). doi:10.1002/sim.6300. https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.6300	20
²¹ 6. Korn, E.L., Graubard, B.I., Midthune, D.: Time-to-Event Analysis of Longitudinal Follow-up of a Survey: Choice of the Time-scale. <i>American Journal of Epidemiology</i> 145 (1), 72–80 (1997). doi:10.1093/oxfordjournals.aje.a009034.	21
²² https://academic.oup.com/aje/article-pdf/145/1/72/150691/145-1-72.pdf	22
²³ 7. Thiébaud, A.C.M., Bénichou, J.: Choice of time-scale in cox's model analysis of epidemiologic cohort data: a simulation study. <i>Statistics in Medicine</i> 23 (24), 3803–3820 (2004). doi:10.1002/sim.2098. https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.2098	23
²⁴ https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.2098	24
²⁵ 8. Pencina, M.J., Larson, M.G., D'Agostino, R.B.: Choice of time scale and its effect on significance of predictors in longitudinal studies. <i>Statistics in Medicine</i> 26 (6), 1343–1359 (2007). doi:10.1002/sim.2699. https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.2699	25
²⁶ https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.2699	26
²⁷ 9. Chalise, P., Chicken, E., McGee, D.: Performance and prediction for varying survival time scales. <i>Communications in Statistics - Simulation and Computation</i> 42 (3), 636–649 (2013). doi:10.1080/03610918.2011.650259. https://doi.org/10.1080/03610918.2011.650259	27
²⁸ 10. Farewell, V.T., Cox, D.R.: A note on multiple time scales in life testing. <i>Journal of the Royal Statistical Society. Series C (Applied Statistics)</i> 28 (1), 73–75 (1979)	28
²⁹ 11. Oakes, D.: Multiple time scales in survival analysis. <i>Lifetime data analysis</i> 1 (1), 7–18 (1995). doi:10.1007/bf00985253	29
³⁰ 12. Duchesne, T., Lawless, J.: Alternative time scales and failure time models. <i>Lifetime data analysis</i> 6 (2), 157–179 (2000). doi:10.1023/a:1009616111968	30
³¹ 13. Cox, D.R.: Regression models and life tables. <i>Journal of the Royal Statistical Society B</i> (34), 187–202 (1972)	31
³² 14. Rutherford, M.J., Crowther, M.J., Lambert, P.C.: The use of restricted cubic splines to approximate complex hazard functions in the analysis of time-to-event data: a simulation study. <i>Journal of Statistical Computation</i>	32
³³	33

- 1 and Simulation **85**(4), 777–793 (2015). doi:10.1080/00949655.2013.845890. 1
- 2 <https://doi.org/10.1080/00949655.2013.845890> 2
- 3 15. Bower, H., Crowther, M.J., Rutherford, M.J., Andersson, T.M.-L., Clements, M., Liu, X.-R., Dickman, P.W., 3
- 4 Lambert, P.C.: Capturing simple and complex time-dependent effects using flexible parametric survival models: 3
- 5 A simulation study. Communications in Statistics - Simulation and Computation **0**(0), 1–17 (2019). 4
- 6 doi:10.1080/03610918.2019.1634201. <https://doi.org/10.1080/03610918.2019.1634201> 4
- 7 16. Syriopoulou, E., Mozumder, S.I., Rutherford, M.J., Lambert, P.C.: Robustness of individual and marginal 5
- 8 model-based estimates: A sensitivity analysis of flexible parametric models. Cancer Epidemiology **58**, 17–24 6
- 9 (2019). doi:10.1016/j.canep.2018.10.017 6
- 10 17. Akaike, H.: A new look at the statistical model identification. IEEE Transactions on Automatic Control **19**(6), 7
- 11 716–723 (1974). doi:10.1109/TAC.1974.1100705 7
- 12 18. Schwarz, G.: Estimating the dimension of a model. Ann. Statist. **6**(2), 461–464 (1978). 8
- 13 doi:10.1214/aos/1176344136 9
- 14 19. Volinsky, C.T., Raftery, A.E.: Bayesian information criterion for censored survival models. Biometrics **56**(1), 10
- 15 256–262 (2000) 10

12 Figures 12

13 13

14 14

15 **Figure 1** Survival proportions, hazard rates and log hazard ratios of the seven data generating 15

16 mechanisms over time t since relapse given relapses occurred at $c_0 = 0.5$, 4 years after 16

17 transplantation. Note that the y -scales differ across scenarios. 17

18 18

19 19

20 20

21 **Figure 2** Bias in survival proportions from fitted models for $x = 0$ for $t = 1, 3, 5$ years since 21

22 relapse given relapse times at $c_0 = 0.5$, 4 years after transplantation. Note that the y -scales differ 22

23 across scenarios. 23

24 24

25 25

26 **Figure 3** Bias in survival proportions from fitted models for $x = 1$ for $t = 1, 3, 5$ years since 26

27 relapse given relapse times at $c_0 = 0.5$, 4 years after transplantation. Note that the y -scales differ 27

28 across scenarios. 28

29 29

30 30

31 **Figure 4** Bias in log hazard ratios from models for $t = 1, 3, 5$ years since relapse given relapse 31

32 times at $c_0 = 0.5$, 4 years after transplantation. Note that the y -scales differ across scenarios. 32

33 33

¹Tables

2

3

Table 1 Data-generating mechanisms for 1000 simulated datasets with 1000 observations with two time-scales: first time-scale, t , and second time-scale, $t + c_0$. *Int* - interaction

DGM	non PH-x on t_1	non PH-x on t_2	Int $t_1 \cdot t_2$	Formula
S_1				$\log(h(t; x)) = -0.5 + 0.1 \log(t) - 0.03 \log(t)^2 - 0.1(t + c_0) + 0.003(t + c_0)^2 - 0.3x$
S_2			x	$\log(h(t; x)) = -0.5 + 0.1 \log(t) - 0.03 \log(t)^2 - 0.1(t + c_0) + 0.003(t + c_0)^2 - 0.3x - 0.1 \log(t) \cdot (t + c_0)$
S_3	x			$\log(h(t; x)) = -0.5 + 0.1 \log(t) - 0.03 \log(t)^2 - 0.1(t + c_0) + 0.003(t + c_0)^2 - 0.3x - 0.2x \log(t)$
S_4	x		x	$\log(h(t; x)) = -0.5 + 0.1 \log(t) - 0.03 \log(t)^2 - 0.1(t + c_0) + 0.003(t + c_0)^2 - 0.3x - 0.2x \log(t) - 0.1 \log(t) \cdot (t + c_0)$
S_5		x		$\log(h(t; x)) = -0.5 + 0.1 \log(t) - 0.03 \log(t)^2 - 0.1(t + c_0) + 0.003(t + c_0)^2 - 0.3x - 0.2x(t + c_0)$
S_6		x	x	$\log(h(t; x)) = -0.5 + 0.1 \log(t) - 0.03 \log(t)^2 - 0.1(t + c_0) + 0.003(t + c_0)^2 - 0.3x - 0.2x(t + c_0) - 0.1 \log(t) \cdot (t + c_0)$
S_7	x	x		$\log(h(t; x)) = -0.5 + 0.1 \log(t) - 0.03 \log(t)^2 - 0.1(t + c_0) + 0.003(t + c_0)^2 - 0.3x - 0.2x \log(t) - 0.2x(t + c_0)$

Table 2 Flexible parametric models with one time-scale used to fit the simulated data with two time-scales. *Int* - interaction.

Fitted model	Corresponds to DGM	non PH-x	Int $x \cdot c_0$	non PH- c_0	Formula
M1	S_1				$\log(h(t; x)) = s_1(\log(t)) + s_2(c_0) + \beta_1 x$
M2	S_2			x	$\log(h(t; x)) = s_1(\log(t)) + s_2(c_0) + \beta_1 x + \psi(s_2(c_0) \cdot \log t)$
M3	S_3	x			$\log(h(t; x)) = s_1(\log(t)) + s_2(c_0) + \beta_1 x + s_3(x \cdot \log t)$
M4	S_4	x		x	$\log(h(t; x)) = s_1(\log(t)) + s_2(c_0) + \beta_1 x + s_3(x \cdot \log t) + \psi(s_2(c_0) \cdot \log t)$
M5	S_5		x		$\log(h(t; x)) = s_1(\log(t)) + s_2(c_0) + \beta_1 x + \gamma(x \cdot s_2(c_0))$
M6	S_6		x	x	$\log(h(t; x)) = s_1(\log(t)) + s_2(c_0) + \beta_1 x + \gamma(x \cdot s_2(c_0)) + \psi(s_2(c_0) \cdot \log t)$
M7	S_7	x	x		$\log(h(t; x)) = s_1(\log(t)) + s_2(c_0) + \beta_1 x + s_3(x \cdot \log t) + \gamma(s_2(c_0))$

Table 3 Percentage of times the fitted models showed lowest AIC or BIC relative to other models when fitted to the same dataset from each data-generating mechanism. Cells in bold represent the highest percentage.

Fitted models	S_1		S_2		S_3		S_4		S_5		S_6		S_7	
	AIC, %	BIC, %	AIC, %	BIC, %	AIC, %	BIC, %	AIC, %	BIC, %	AIC, %	BIC, %	AIC, %	BIC, %	AIC, %	BIC, %
M1	65.9	99.8	0.4	27.2	8.4	73.1	0	6.3	0	6.4	0	25.6	0	0
M2	10.8	0	76	72.7	1.3	0	6.5	51.1	0	0	2.7	26.7	0	0
M3	9.9	0.2	0	0	67.6	26.9	0	3.2	12.7	73.3	0.4	19.3	12.8	78.8
M4	1.1	0	13	0	11.8	0	92.2	39.4	3	0	49	14.7	3.7	0.1
M5	8.6	0	0.1	0	0.7	0	0	0	0.5	2.9	0.5	7.9	0	0
M6	1.7	0	10.5	0.1	0.1	0	1.3	0	0.1	0	28	1.4	0	0
M7	2	0	0	0	10.1	0	0	0	83.7	17.4	19.4	4.43	83.5	21.1

14

Additional Files

15

Additional file 1

16A

HTML file containing all the results from the data analyses with additional figures and tables.

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

Figures

Figure 1

Survival proportions, hazard rates and log hazard ratios of the seven data generating mechanisms over time t since relapse given relapses occurred at $c_0 = 0.5; 4$ years after transplantation. Note that the y-scales differ across scenarios.

Bias in $S(t)$ for $x = 0$

□ M1 ■ M2 ○ M3 ● M4 △ M5 ▲ M6 ◇ M7

Figure 2

Bias in survival proportions from fitted models for $x = 0$ for $t = 1; 3; 5$ years since relapse given relapse times at $c_0 = 0.5; 4$ years after transplantation. Note that the y-scales differ across scenarios.

Bias in $S(t)$ for $x = 1$

□ M1 ■ M2 ○ M3 ● M4 △ M5 ▲ M6 ◇ M7

Figure 3

Bias in survival proportions from fitted models for $x = 1$ for $t = 1; 3; 5$ years since relapse given relapse times at $c_0 = 0.5; 4$ years after transplantation. Note that the y-scales differ across scenarios.

Figure 4

Bias in log hazard ratios from models for $t = 1; 3; 5$ years since relapse given relapse times at $c_0 = 0.5; 4$ years after transplantation. Note that the y-scales differ across scenarios.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementarymaterial.zip](#)