ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Journal of the Royal Statistical Society
Applied Statistics
Series C

*Appl. Statist.* (2019)
**68**, *Part* 4, *pp.* 985–1005

# An instrumental variable procedure for estimating Cox models with non-proportional hazards in the presence of unmeasured confounding

Pablo Martínez-Camblor, Todd A. MacKenzie and Douglas O. Staiger

*Dartmouth College, Hanover, USA*

Phillip P. Goodney

*Dartmouth College, Hanover, and Dartmouth–Hitchcock Medical Center, Lebanon, USA*

and A. James O'Malley

*Dartmouth College, Hanover, USA*

[Received May 2018. Revised January 2019]

**Summary.** Two-stage instrumental variable methods are commonly used for estimating average causal effects in the presence of an unmeasured confounder. In the context of the proportional hazard Cox regression models, this problem has recently received attention with several methods being proposed. Previously, we developed an improved estimator under the incumbent two-stage residual inclusion procedure called '2SRI' by adding a Gaussian frailty in the second stage. We now consider the more complex situation in which the treatment and the unmeasured confounders can have time varying effects, illustrating the method with the case of a step function with one prespecified change point. We prove that, in situations where the effects of the unmeasured confounder or the treatment change during the follow-up, the first stage of the 2SRI algorithm induces a frailty with time varying coefficients in the second stage, which enables incumbent methods and our previously developed procedure to be improved on. A Monte Carlo simulation study demonstrates the superior performance of the proposed extension of 2SRI that we develop. We apply the new procedure to estimate the effect of endarterectomy *versus* carotid artery stenting on the time to death of patients suffering from carotid artery disease by using linked vascular quality initiative registry–Medicare data.

*Keywords*: Causal inference; Control function; Cox regression model; Time-dependent coefficients; Time-modified confounding

## 1. Introduction

The semiparametric Cox proportional hazard model (Cox, 1972) provides a useful approach for understanding time-to-event processes. It assumes that each covariate effect on the hazard is the same through the entire study period. The hazard function, which is the instantaneous risk that a subject at risk till $t$ suffers the event at time $t$, is given by

$$\lambda(t|X, U) = \lambda_0(t) \exp(\beta_X X + \beta_U U),$$

*Address for correspondence*: Pablo Martínez-Camblor, Dartmouth Institute for Health Policy and Clinical Practice, Geisel School of Medicine, Dartmouth College, Suite 309, 7 Lebanon Street, Hinman Box 7251, Hanover, NH 03755, USA.
E-mail: Pablo.Martinez.Camblor@Dartmouth.edu

 0035–9254/19/68985

where $X$ and $U$ are covariates and $\lambda_0(\cdot)$ is the baseline hazard function. When $U$ is observed and adjusted for, the model accounts for the possible relationships between the covariates $X$ and $U$ and between the covariate $U$ and the time-to-event $T$. Assuming that the model is correctly specified, $\beta_X$ is the *log-causal-hazard ratio* of $X$ on the studied event and measures the average change in the risk produced by $X$ when $U$ is fixed. One relevant handicap of Cox regression models is the loss of causal interpretation if $U$ is not included in the model, even when there is no interaction between $X$ and $U$ (Aalen *et al.*, 2015a; Martinussen and Vansteelandt, 2013), i.e., if in the above model $\beta_U \neq 0$, then estimating $\lambda(t|X) = \lambda_0^*(t)\exp(\beta_X^* X)$ yields estimates $\beta_X^* \neq \beta_X$ even when $X$ and $U$ are independent. Note that the subjects who are at risk beyond each particular event time are a subset of the individuals who have not previously failed. Therefore, the balance in the distribution of the potential confounder between treatment groups is lost by the implicit conditioning. In this context, instrumental variable (IV) methods that are used to adjust for unmeasured confounding frequently report biased results (Wan *et al.*, 2015). Martínez-Camblor *et al.* (2019) proposed to add an individual (univariate) frailty term in the second stage of the two-stage residual inclusion procedure called '2SRI' (Terza *et al.*, 2008) for dealing with this bias. The frailty term in the second stage accounts for the noise in the residual from the first stage. Under general conditions, the new two-stage residual inclusion–frailty procedure (called '2SRI-F') obtains unbiased estimates of the log-hazard ratio in the presence of unmeasured confounding. Monte Carlo simulations revealed substantial bias reductions even when some theoretical assumptions were not satisfied. Despite detractors from a non-parametric causal perspective (Hernán, 2010), the proportional hazard assumption and the hazard ratio are a useful way of summarizing a (potentially) complex reality. However, in research fields like biomedicine, it is not unusual that patients experience varying risk over the follow-up period following procedural interventions. For example, risks of mortality immediately after surgery from many procedures are transiently higher as patients recover from their procedure and may be subject to complications (see, for instance, Goodney *et al.* (2011)). After patients have recovered from surgery, their risks of mortality return to their baseline rate, perhaps even a lower rate if their surgery mitigated a risk to their survival. Fortunately, one of the Cox model's strengths is its ability to encompass different scenarios such as time-dependent covariates and/or coefficients.

Most of the circumstances that are associated with time-dependent events can change during follow-up. Many references have dealt with this type of problem. For instance, Robins (1986) considered the case where there is a time varying cause of disease that brings about changes in a time varying treatment. Platt *et al.* (2009) studied the case where the relationship between the treatment and confounding changes across follow-up (time-modified confounding). Gran *et al.* (2018) dealt with the problem of estimating the treatment effect on the treated under time-dependent confounding. Recently, Burne and Abrahamowicz (2019) proposed and compared alternative methods for controlling for a time varying unmeasured confounder in large data sets by using additional information from small validation samples.

To estimate a causal effect in a non-linear model in the presence of unmeasured confounding, the most commonly used two-stage procedures are two-stage predictor substitution (Greene and Zhang, 2003) and the 2SRI (also called *control function*) methods. Whereas both methods share a first stage in which the relationship between the treatment and the IV is estimated by using any consistent estimation technique, the second stage differs: in the two-stage predictor substitution the observed exposure is replaced by its prediction obtained in the first stage; 2SRI includes in the second stage the residuals that are computed in the first stage as an additional covariate. To the best of our knowledge, no published work has studied the performance of these procedures in Cox regression models with time-dependent treatment effects.

In this paper, we deal with the problem of estimating the causal hazard ratio $\exp(\beta_X)$ in the

presence of unmeasured covariates when it is time dependent, denoted as $\beta_X(t)$. In particular, we consider the simple case in which $\beta_X(t) = \beta_{X,1}\, I_{[0,t^*)}(t) + \beta_{X,2}\, I_{[t^*,\infty)}(t)$, with $t^*$ a fixed (known) point of time and $I_A$ the indicator function ($I_A(t)$ takes the value 1 if $t \in A$ and 0 otherwise). In addition, the behaviour of the 2SRI-F procedure when the unmeasured confounding effect changes at this point of time ($\beta_U(t) = \beta_{U,1}\, I_{[0,t^*)}(t) + \beta_{U,2}\, I_{[t^*,\infty)}(t)$), i.e. the presence of time-modified confounding is also explored. The two-stage framework is presented in Section 2. In Section 3, we prove that, when the unmeasured covariates are time modified, the procedure 2SRI induces time-dependent frailty in the second stage. We propose to estimate the parameters of the resulting model by maximizing the *integrated partial likelihood function* (Therneau *et al.*, 2003). In Section 4, we study the behaviour of the 2SRI-F algorithm on finite samples via Monte Carlo simulations. In Section 5, we estimate the therapeutic effect on all-cause mortality of the received treatment (*endarterectomy versus carotid stenting*) on carotid artery disease patients from the 'vascular quality initiative' registry–Medicare linked, the problem which motivated this research. The initiative is a national registry of patients treated for vascular disease, where the benefits of treatment—such as removing blockages from the carotid arteries, which supply blood to the brain—are measured over time. This treatment, which includes procedures known as carotid endarterectomy and carotid stenting, is proffered by a surgical intervention, which carries an up-front risk during the acute period of complications from surgery. This risk is then observed to equalize and eventually to be surpassed by a time-dependent treatment effect, measured in terms of long-term stroke risk reduction. We allow both the treatment effect and the unmeasured confounding to change between the acute period (the first 30 days) and the long-term follow-up (10 years). The proposed model and estimation approach enable a better understanding of the treatment effect than the simple application of 2SRI-F. Finally, in Section 6, we present our main conclusions, whereas Appendix A provides some additional figures and some R code that was used for implementing the proposed procedures.

The programs that were used to analyse the data can be obtained from

`https://rss.onlinelibrary.wiley.com/hub/journal/14679876/series-c-datasets`

## 2. The two-stage instrumental variable framework

In the two-stage modelling framework (Terza *et al.*, 2008) it is typically assumed that both the outcome studied, $T$, and the treatment assignment, $X$, share the same unmeasured covariates. In the linear context, this is not a restrictive assumption because of the greater affinity of procedures with explicit additive error terms for dealing with independent unmeasured covariates (i.e. variables that predict the outcome but not the treatment assignment). However, this is a concern in the Cox proportional hazard model. Indeed, the causal interpretation of the hazard ratio is compromised in the presence of unmeasured covariates that are unrelated to the treatment (Aalen *et al.*, 2015a), which manifests as an individual random effect in the survival model—known as a frailty in survival analysis parlance. In addition, Martínez-Camblor *et al.* (2019) proved that, in the control function method, a first-stage including unmeasured covariates related to the treatment assignment but unrelated to the main outcome induces a frailty in the second-stage survival model. Under the causal proportional hazard model given by

$$\lambda(t|X,U) = \lambda_0(t)\exp(\beta_X X + \beta_U U), \tag{1}$$

$U$ depicts the unmeasured confounding, $\beta_U$ is the log-hazard (coefficient) of the unmeasured

covariates, $X$ is the received treatment, $\beta_X$ is our target and $\lambda_0(\cdot)$ is the baseline hazard function. The treatment received is the result of the selective process depicted by the equation

$$X = \alpha_0 + \alpha_W W + \alpha_U U + \epsilon, \tag{2}$$

with $\epsilon$ white noise and $W$ a measured random variable satisfying the following conditions.

*Condition 1*. $(W \not\perp\!\!\!\perp X)|U$.

*Condition 2*. $(W \perp\!\!\!\perp T)|X, U$.

*Condition 3*. $W \perp\!\!\!\perp U$ (randomization assumption).

Then $W$ can be considered an IV. The strength of the instrument reflects the strength of the relationship between $W$ and $X$. In a randomized trial, assigned treatment is an instrument. Assumptions 1–3 can be reformulated and combined with the stable unit treatment value assumption and the monotonicity assumption (Hernán and Robins, 2006) between the treatment and the IV to complete the conditions under which the IV $W$ identifies the effect of $X$ non-parametrically (i.e. without relying on equations (1) and (2)). Common instruments include prior institutional affinity for using a particular procedure, geographic region of residence, an individual's differential access to certain treatments and an individual's genes (also known as Mendelian randomization (Thanasassoulis and O'Donnell, 2009)). Although our method requires that $W$ be independent of $U$ and associated with $X$, but does not require that $W$ have a causal effect on $X$, the key point is that, in the first-stage model for the treatment assignment, $W$ captures information about confounders: it is not required that $W$ be causal. Fig. 1 depicts the directed acyclic graph (Pearl, 1995), with $W$ a non-causal and $W^*$ a causal IV (Swanson and Hernán, 2018). Pathways connecting covariates with the outcome at different moments (the grey arrows) are the main handicap of the IV in the time-to-event context. Of course, a vector of measured covariates can be easily accommodated within this framework. We do not explicitly depict any measured covariate herein. Without loss of generality and to gain simplicity in the expressions reported, we assume that $U$ and $W$ are centred at 0.

In the first stage of procedure 2SRI, we estimate the residuals

$$R = X - (\alpha_0 + \alpha_W W) = \alpha_U U + \epsilon,$$

that are included in the estimation of the survival model in the second stage:

$$\begin{aligned}
\mathcal{P}\{T > t|X = x, U = \alpha_U^{-1}(r-\epsilon)\} &= \mathbb{E}_\epsilon\left[\exp\left[-\int_0^t \lambda_0(s)\exp\{\beta_X x + \beta_U \alpha_U^{-1}(r-\epsilon)\}\mathrm{d}s\right]\right] \\
&= \mathbb{E}_\epsilon\left[\exp\left\{-\int_0^t z(\epsilon)\lambda_0(s)\exp(\beta_X x + \beta_R r)\mathrm{d}s\right\}\right] \\
&= \mathcal{P}(T > t|X = x, R = r),
\end{aligned}$$

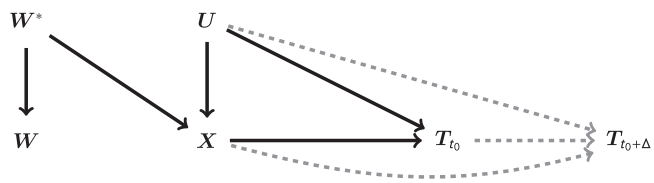


**Fig. 1.** Directed acyclic graph showing a causal $W^*$ and a non-causal $W$ IV: $U$ stands for an unmeasured confounder

where $z(\epsilon) = \exp(-\beta_U \alpha_U^{-1} \epsilon)$ and $\beta_R = \beta_U \alpha_U^{-1}$, i.e. a Cox regression model with frailty term is determined as the statistical model to estimate in the second step. To identify the model, we assume that the frailties follow a parametric distribution. Herein, we suppose that the frailty term $z(\epsilon)$ that is included in the final Cox regression model follows a log-normal distribution $\mathcal{N}(0, \sigma_\epsilon)$.

In the current work we consider the more complex situation in which both the treatment and the unmeasured confounder effects are allowed to change with time, i.e. we assume the causal model

$$\lambda(t|X, U) = \lambda_0(t) \exp\{\beta_X(t) X + \beta_U(t) U\}, \tag{3}$$

where $U$ depicts an unmeasured confounder, $\beta_U(\cdot)$ is the log-hazard (coefficient) of the unmeasured covariates and $X$ is the received treatment. Our goal is to estimate the treatment effect $\beta_X(\cdot)$, assuming the assignment model (2), the hypothesis 1–3 and the specified frailty distribution.

## 3. Frailties in the 2SRI algorithm

The concept of a *random effect* is translated to time-to-event models as a *frailty* (Hougaard, 1995), i.e. a *frailty Cox proportional hazard regression model* is a standard Cox regression model which includes a random effect (frailty). The frailty reflects the susceptibility of a subject (univariate frailty) or a group (multivariate frailty) suffering the event studied. There is an increasing number of references dealing with both theoretical and practical aspects of this topic. Readers who are interested in technical aspects of frailties are referred to Wienke (2010) and the references therein. Several real world situations in which the frailty term is useful for understanding the time-to-event process were reported by Aalen *et al.* (2015b).

The previous framework (Section 2) directly implies that

$$\begin{aligned} \mathcal{P}(T > t|X = x, R = r) &= \mathbb{E}_\epsilon \left[ \exp\left[ -\int_0^t \lambda_0(s) \exp\{\beta_X(s) x + \beta_U(s) \alpha_U^{-1}(r - \epsilon)\} \mathrm{d}s \right] \right] \\ &= \mathbb{E}_\epsilon \left[ \exp\left[ -\int_0^t z(s, \epsilon) \lambda_0(s) \exp\{\beta_X(s) x + \beta_R(s) r\} \mathrm{d}s \right] \right], \end{aligned} \tag{4}$$

where $r = x - \alpha_0 - \alpha_W w$, $z(s, \epsilon) = \exp\{-\beta_U(s) \alpha_U^{-1} \epsilon\}$ and $\beta_R(s) = \beta_U(s) \alpha_U^{-1}$, i.e. the previous Cox proportional hazard model with frailty term enables the targeted parameter $\beta_X(s)$ to be estimated. Note that, under assumptions 1–3 and for adequately distributed frailties, all components of the model can be estimated from observed data.

In general, it is easy to control for measured confounders. For simplicity, we do not include the term $Z$ in the method's development. However, in the case that a measured confounder $Z$ is included in both the survival and the assignment models with coefficients $\beta_Z$ and $\alpha_Z$ respectively, it is not possible to assure that the estimate of $R = X - (\alpha_0 + \alpha_W W + \alpha_Z Z)$ is unbiased without assuming $Z \perp\!\!\!\perp U$; we do not have an unbiased estimator for $\alpha_Z$ and, therefore, this bias transmits to the final estimation of $\beta_Z(\cdot)$. When there are interactions between the treatment and measured confounders, the procedure can control for this effect and report consistent estimates for the treatment effect but cannot report the correct effect of the interaction itself. An alternative procedure for estimating the interaction with measured confounders is to estimate separate models.

Particular families of frailties have been considered in the survival analysis literature. For instance, Unkel *et al.* (2014) considered time-dependent frailties in the context of infectious

disease. In the problem currently considered, the identifiability of both the frailty time-dependent coefficient $-\beta_U(\cdot)\alpha_U^{-1}$ and the variance of the random component, $\epsilon$, is well studied. However, since $-\beta_U(\cdot)\alpha_U^{-1}$ scales the frailty distribution, we can fix the variance of $\epsilon$ to be equal to 1.

Although frailties are helpful towards obtaining good estimates for the parameters of interest, they complicate the estimation of the predicted values for new subjects. The effect of the treatment on the difference between the survival curves depends on the length of follow-up, $t$, and on the value of the unmeasured confounder. This effect is smallest for the most extreme values of the survival probabilities, i.e. at the greatest and lowest $U$-values (see Fig. 6 in Appendix A.1).

### 3.1. Parameter estimation

In practice, we often have the sample $\{(y_i, \Delta_i)\}_{i=1}^N$, where $y_i = \min\{t_i, c_i\}$ and $\Delta_i = I_{(-\infty, c_i]}(t_i)$, and $c_i$ and $t_i$ are the censoring and the survival times for the $i$th subject ($1 \leqslant i \leqslant N$) respectively. For each subject we also know the IV values, $\{w_i\}_{i=1}^N$, and the treatment received, $\{x_i\}_{i=1}^N$. To guarantee asymptotic convergence for the standard Cox regression model estimator, a sufficient condition is that the censoring times $C$ satisfy $(C \perp\!\!\!\perp Y)|X, U, W$. The second-stage goal is then to fit the model

$$\lambda(t|X, R, \delta) = \lambda_0(t)\exp\{\beta_X(t)X + \beta_R(t)R + \beta_F(t)\delta\},$$
$$\delta \sim \mathcal{N}(0, 1),$$

where $\beta_F(t)$ is the frailty time-dependent coefficient. Assuming that, for a fixed point of time $t^*$, we have that

$$\begin{aligned}\mathbf{B}(\mathbf{s}) &= (\beta_X(s), \beta_R(s), \beta_F(s))\\ &= (\beta_{X,1}, \beta_{R,1}, \beta_{F,1})\, I_{[0,t^*]}(s) + (\beta_{X,2}, \beta_{R,2}, \beta_{F,2})\, I_{(t^*,\infty)}(s) = \mathbf{B}_1 + \mathbf{B}_2.\end{aligned}$$

We can then define the usual Cox partial likelihood function and integrate out the random effects to create the integrated partial likelihood (Therneau *et al.*, 2003):

$$\begin{aligned}\mathbf{i}(\mathbf{B}_1, \mathbf{B}_2) &= \sum_{i=1}^n \int\int \Delta_i \left[\eta_{1,i} + \eta_{2,i} - \log\left\{\sum_{j=1}^n Y_j(t)\exp(\eta_{1,j} + \eta_{2,i})\right\}\right]\varphi(\delta_i)\mathrm{d}\delta_i\,\mathrm{d}t\\ &= \sum_{i=1}^n \int \left(\int_0^{t^*} \Delta_i \left[\eta_{1,i} - \log\left\{\sum_{j=1}^n Y_{1,j}(t)\exp(\eta_{1,j})\right\}\right]\mathrm{d}t\right.\\ &\quad \left. + \int_{t^*}^{\infty} \Delta_i \left[\eta_{2,i} - \log\left\{\sum_{j=1}^n Y_j(t)\exp(\eta_{2,j})\right\}\right]\mathrm{d}t\right)\varphi(\delta_i)\,\mathrm{d}\delta_i\end{aligned}$$

where $\eta_{k,l} = \beta_{X,k}x_l + \beta_{R,k}r_l + \beta_{F,k}\delta_l$ ($k = 1, 2$ and $l = 1, \ldots, n$), $Y_l(t)$ describes the risk set ($Y_l(t) = 1$ if the $l$th subject is still under observation (without the event) at time $t$ and $Y_l(t) = 0$ otherwise) and $\varphi(\cdot)$ is the density function of a standard normal distribution. Ripatti and Palmgren (2000) showed the connection between the Gaussian frailty Cox model and a penalized likelihood procedure and used this result to derive an estimation procedure. Next, we propose an algorithm for approximating the solution to the above equation.

The algorithm proposed estimates the first-period frailties and then solves for the implied second-period frailty term. The algorithm is as follows.

*Step 1*: estimate the treatment assignment model $X = \hat{\alpha}_W W$ and compute $\hat{R} = X - \hat{\alpha}_W W$ (first stage).

*Step 2*: compute the first-period Cox regression model with an individual frailty term and $\hat{R}$ as covariate. Times are censored at time $t^*$ if the event studied has not occurred by then.

*Step 3*:

(a) from the first-period model, compute $\hat{\beta}^{\mathrm{IV}}_{X,1}$ (the estimator of $\beta_{X,1}$), and $\hat{R}_1 = \hat{R} + \hat{\beta}^{-1}_{R,1}\hat{\phi}_1$, where $\hat{\phi}_1$ are the log-frailty estimates provided in the estimation of the Cox regression model with individual frailty term (second stage–first period);

(b) estimate a second-period Cox regression model with individual Gaussian frailties including $X$, and $\hat{R}_1$ as covariates; the resulting $\hat{\beta}^{\mathrm{IV}}_{X,2}$ is the estimator for $\beta_{X,2}$ (second stage—second period).

Note that, under model (2), assuming without loss of generality that $\alpha_0 = 0$, conditions 1 and 2 guarantee that $\hat{R}$ is an unbiased estimator for $R = \alpha_U U + \epsilon$. Assuming that $\epsilon$ is normally distributed white noise, theoretical developments in Martínez-Camblor *et al.* (2019) guarantee that, if the censoring time and the time to event given $X$, $U$ and $W$ are independent, $\sqrt{n}\{\hat{\beta}^{\mathrm{IV}}_{U,1} - \beta_{U,1}\}$ converges weakly to a normal distribution. Therefore, under these assumptions, step 3(a) provides a consistent estimator for $\beta_{X,1}$.

We show in equation (4) that $\beta_{U,1}U = \beta_{R,1}R + \phi$, where $\phi$ $(= \sigma_\phi \delta)$ is the frailty that is actually present in the estimation of the first-period model in step 2 above ($\lambda(t|X = x, R = r) = \lambda_0(t)\exp(\beta_{X,1}x + \beta_{R,1}u + \phi)$ with $t \in [0, t^*)$ and $\phi$ follows an $\mathcal{N}(0, \sigma_\phi)$ distribution). Since $\beta_{R,1} = \beta_{U,1}\alpha_U^{-1}$, then $\alpha_U U = R + \beta^{-1}_{R,1}\phi = R_1$. Arguing as in equation (4), we have for $t \in [t^*, \infty)$ that

$$\begin{aligned}
&\mathcal{P}(T > t|X = x, R_1 = r_1)\\
&\quad = \mathcal{P}(T > t|X = x, U = \alpha_U^{-1} r_1)\\
&\quad = \exp\left\{ -\int_0^{t^*} \lambda_0(s)\exp(\beta_{X,1}x + \beta_{U,1}\alpha_U^{-1}r_1)\mathrm{d}s + \int_{t^*}^{t} \lambda_0(s)\exp(\beta_{X,2}x + \beta_{U,2}\alpha_U^{-1}r_1)\mathrm{d}s \right\}.
\end{aligned}$$

We cannot guarantee that the individual values of $\hat{R}_1$ (the empirical estimator of $R_1$) converge to the individual values of $R_1$. However, assuming that $\hat{\phi} \sim \mathcal{N}(\phi, \sigma_\phi)$ (this assumption is true if $\phi$ is a Gaussian frailty term) and that $\hat{R} \sim \mathcal{N}(R, \sigma_R)$ (the usual assumption for making inference in linear models), then $\hat{R}_1 = R_1 + \tau$, where $\tau$ is a normally distributed random variable. In general, $\hat{\phi}_1 \perp\!\!\!\perp \hat{R}_1$. Therefore,

$$\begin{aligned}
\mathcal{P}(T > t|X = x, \hat{R}_1 = \hat{r}_1) &= \mathcal{P}(T > t|X = x, R_1 = r_1 + \tau)\\
&= \mathbb{E}_\tau\left[\exp\left[ -\int_0^t z^*(s,\tau)\lambda_0(s)\exp\{\beta_X(s)x - \beta_U(s)\alpha_U^{-1}r_1\}\mathrm{d}s \right]\right],
\end{aligned}$$

where $z^*(s,\tau) = \exp\{-\beta_U(s)\alpha_U^{-1}\tau\}$, i.e. the Cox proportional hazard model with Gaussian frailty term that is proposed in the second stage for the second-period estimates the parameter $\beta_{X,2}$. Arguing as in Martínez-Camblor *et al.* (2019), under correctly assumed parametric assumptions for the error terms, we have that the estimator is consistent and asymptotically normally distributed.

Although some asymptotic theory has been developed for different frailty structures and estimation procedures (see, for instance, Murphy (1995) and Kosorok and Lee (2004)), in the univariate case computational approximations or other idiosyncrasies in available software packages mean that, even under optimal situations, some bias remains. Barker and Henderson (2005) pointed out that, for univariate gamma frailties, the usual expectation–maximization approach leads to finite sample underestimation of the frailty variance, with the corresponding regression parameters also being underestimated as a result. Fig. 7 (in Appendix A.1) illustrates this issue.

### 3.2. Alternative 2SRI-F extensions to account for time varying treatment effects

The theory underlying equation (3) suggests that the above algorithm is the correct extension

of the 2SRI-F estimation procedure to the time-dependent treatment effect set-up. However, although direct application of 2SRI-F to the first period is clearly appropriate, other natural alternatives can be applied for estimating the coefficient for treatment in the second period. Here, we briefly present two alternatives. These reflect a practitioner who

(a) naively treats survival across the two time periods as unrelated problems or
(b) naively directly applies the 2SRI-F algorithm (Martínez-Camblor *et al.*, 2019) by including the first-stage residual and a single frailty both with fixed coefficients across both time periods.

By evaluating the relative performance of these two alternative procedures that could unwittingly be assumed by practitioners to be the appropriate generalization of 2SRI-F, this section of the paper emphasizes an important contribution that is distinct from prior work (Martínez-Camblor *et al.*, 2019).

The first alternative approach, called '2SRI-$F_{[1]}$', directly applies 2SRI-F to the second time period. By using this algorithm, the estimated frailties in the first and second periods may bear no relationship to one another. The second-period frailties are computed by using only the subset of subjects surviving to the beginning of the second period. Because survivor subsets are selective, the distribution of the frailties for the survivors to the second period is likely to be quite different from that of the full distribution, which we conjecture is likely to confound the frailty term strongly with the treatment effect in the second period and to result in substantial bias.

In the second alternative approach, we conjecture that not allowing the residual and frailty terms to be scaled differently in the second period will fail to capture any change in the true effect of the unmeasured confounder on survival. This claim stems from the following argument. The estimated treatment selection mechanism in the first time period has the form $\beta_{U,1}U = \beta_{R,1}R + \epsilon$, from which the unmeasured confounder $U$ is estimated. We include the estimates of $U$, the first-stage residual, as a covariate in the second period to control approximately for $U$ whose coefficient is fixed to 1; this is only supported by the model if $\beta_{U,1} = \beta_{U,2}$. In addition, under this second alternative procedure, called '2SRI-$F_{[2]}$', the additional frailty that is used to account for estimation-induced random error is excluded altogether. In short, from the Cox regression model with Gaussian frailty,

$$\lambda(t|X, R) = \mathbb{E}_{\epsilon}[\lambda_0(t)\exp(\beta_{X,1}X + \beta_{R,1}R + \epsilon)] \qquad t \in [0, t^*),$$

we estimate $\beta_{X,1}$. Then, from the Cox regression model,

$$\lambda(t|X, R) = \lambda_0(t)\exp(\beta_{X,2}X + \beta_{R,1}R + \epsilon) \qquad t \in [t^*, \infty),$$

we estimate $\beta_{X,2}$. Note that $\epsilon$ are the frailties that are computed in the first period. A critical problem is that, without the additional frailty, the individual frailties that are estimated in the second stage have a sample error which depends on the sample size and, therefore, are not asymptotically null.

## 4. Monte Carlo simulations

We simulate 1000 values (the sample size) for three random variables ($U, \epsilon$ and $W$) following independent standard normal distributions. Then, the continuous treatment assignment or propensity is computed from the model $X = \alpha_W W + U + 2\epsilon$ with $\alpha_W = 1$ (the final variability of the random term in the treatment assignment model is 4). Survival times are generated from a distribution with hazard function $\lambda(t|X, U) = \frac{1}{2}t\exp[\{\beta_{X,1}I_{[0,3)}(t) + \beta_{X,2}I_{[3,\infty)}(t)\}X + \{\beta_{U,1}I_{[0,3)}(t) + \beta_{U,2}I_{[3,\infty)}(t)\}U]$. Censoring times are independently generated from a gamma

model. Observed times above 10 were censored at this point. Censoring distribution parameters are selected to obtain a final censorship average between 15% and 30%. Fig. 2 depicts the median of the 2000 simulated quadratic errors, $E^2 = (\hat{\beta}_{X,1} - \beta_{X,1})^2 + (\hat{\beta}_{X,2} - \beta_{X,2})^2$, for the real (including unmeasured covariates), naive (omitting the presence of unmeasured confounding), 2SRI (two-stage residual inclusion letting the residuals effects change with time) and 2SRI-F (two-stage residual inclusion with Gaussian frailty) procedures for various values of $\beta_{X,1}$ (the $x$-axis), $\beta_{X,2}$ and $\beta_{U,2}$ and $\beta_{U,1} = \log(\frac{1}{3})$. The case $\beta_{U,2} = \log(\frac{1}{3})$ represents the case when the effect of the unmeasured confounder does not change with time. As already mentioned, to implement procedure 2SRI-F we estimate the first-period frailties and they are included as a covariate in the second period with an additional frailty term (full details about the implementation are reported in Appendix A). The best scenario for using the procedure proposed is the presence of strong effects of the unmeasured covariates in both periods (i.e. when $\beta_{U,2}$ is $\log(\frac{1}{3})$ or $\log(3)$) and likewise for the treatment effects ($\beta_{X,1} = \log(\frac{1}{3})$ and $\beta_{X,2} = \log(3)$). The results that were obtained by the algorithms 2SRI-$F_{[1]}$ and 2SRI-$F_{[2]}$ are also depicted in Fig. 8 (in Appendix A.1).

These Monte Carlo simulations reveal that the proposed implementation for procedure 2SRI-F performs well for continuous treatments; it improves the median quadratic error $E^2$ obtained for 2SRI in all situations considered, especially in the most extreme cases (strongest changes in the coefficients).

In general, the two considered alternative extensions of 2SRI-F to the time varying case, 2SRI-$F_{[1]}$ and 2SRI-$F_{[2]}$, produce worse results than the procedure that we developed, which controls for the unmeasured confounder in all the situations considered. The results for 2SRI-$F_{[1]}$ are, however, still better than those for 2SRI-$F_{[2]}$, which produces erratic outcomes (see Fig. 8 in Appendix A.1). Therefore, the procedure that is developed in this paper provides a clear enhancement of the procedure that we previously developed by appropriately accounting for the change in the effect of the unmeasured confounder and the additional random noise that is induced by the inclusion of the frailty estimated in the first time period as a covariate.

We finally consider a binary treatment. Retaining the previous notation and distributions, the treatment assignment is given by the model $X = I_{[0,\infty)}(W + U + \epsilon)$ or, equivalently, $X = W + U + \epsilon + \phi$ where $\phi = W + U + \epsilon - I_{[0,\infty)}(W + U + \epsilon)$. The quantity $\phi$ is obviously not normally distributed (see Fig. 3 (d)) and is related to $W$, implying that assumption 3 is violated and that the operating characteristics of the procedure will not be as optimistic as when $X$ is continuous. We considered the following parameter settings: $\beta_{U,1} = \log(\frac{1}{3})$, $\beta_{U,2} = \log(3)$ all with $\beta_{X,2} = \log(3)$. The median behaviour in the quadratic error, $E^2$, is similar to that obtained for the continuous treatment case although the scale of the error is larger (Fig. 3(a)). The biases of the frailty in both the first and the second period treatment seem to be smaller than in the continuous treatment case. In general, it is noteworthy that 2SRI-F controls the bias of $\beta_{X,1}$ by the largest amount relative to other procedures at the extreme ends of the range of $X$ (Fig. 3(b)). This finding suggests that substantial unmeasured confounding must be present for the new estimator that accounts for treatment effect heterogeneity to improve substantially on the naive adaptation of 2SRI-F and other procedures.

## 5. Real world data application

We apply 2SRI-F to nationwide data from the vascular quality initiative (`http://www.vascularqualityinitiative.org`), on patients who were diagnosed with carotid artery disease (carotid stenosis). These data contain comprehensive information on all patients suffering from carotid stenosis and are continually updated over time. They facilitate determination of the best procedure or treatment approach to use on average and to determine which types of patient

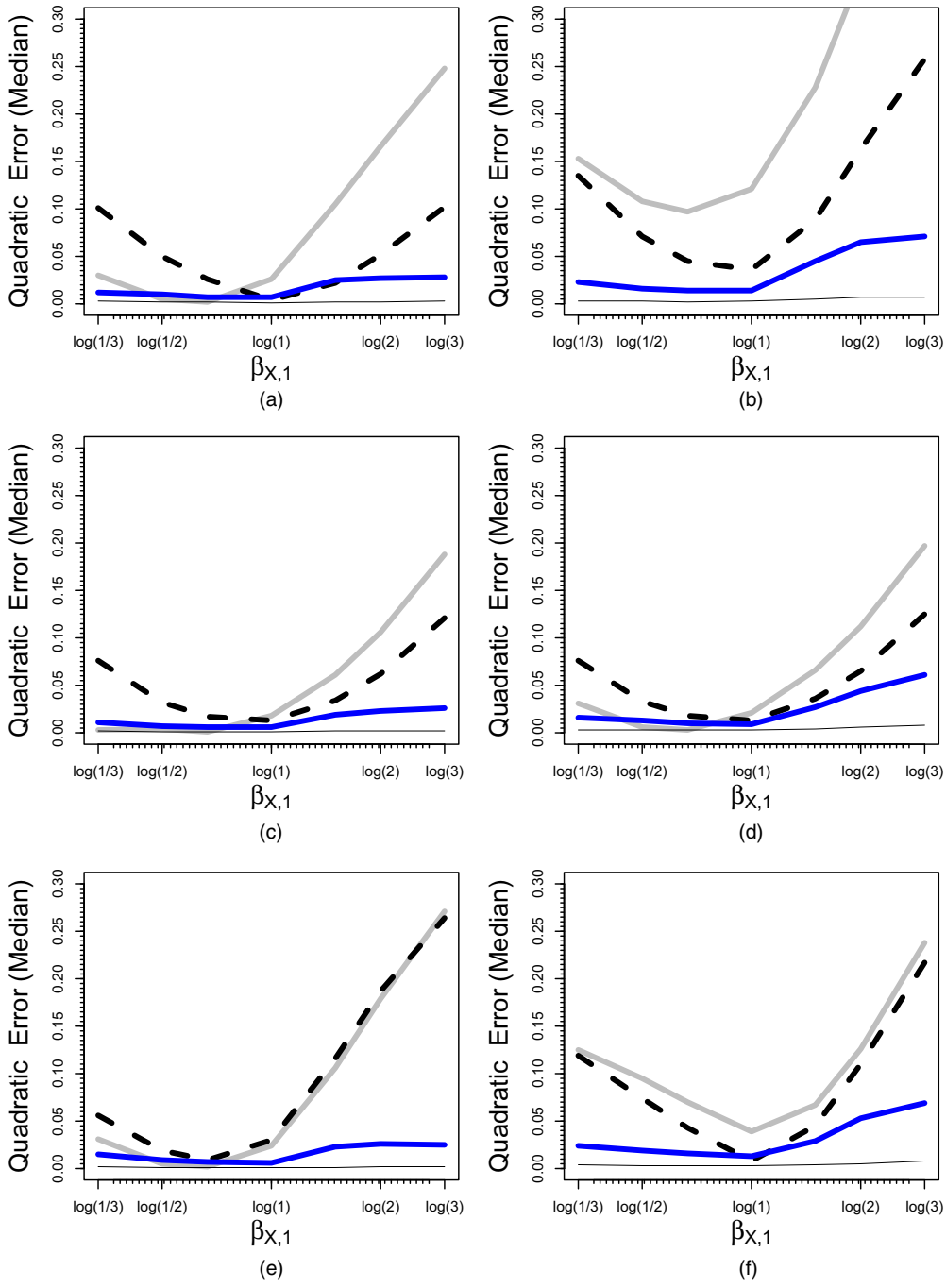


**Fig. 2.** Median of the simulated quadratic errors $E^2 = (\hat{\beta}_{X,1} - \beta_{X,1})^2 + (\hat{\beta}_{X,2} - \beta_{X,2})^2$ for the survival and the assignment models described in the text (they are based on 2000 simulated data sets for the four models; the sample size is 1000; $\beta_{U,1} = \log(\frac{1}{3})$ (———, real; ———, naive; — —, 2SRI; ———, 2SRI-F): (a) $\beta_{X,2} = \log(1)$, $1\beta_{U,2} = \log(\frac{1}{3})$; (b) $\beta_{X,2} = \log(3)$, $\beta_{U,2} = \log(\frac{1}{3})$; (c) $\beta_{X,2} = \log(1)$, $\beta_{U,2} = \log(1)$; (d) $\beta_{X,2} = \log(3)$, $\beta_{U,2} = \log(1)$; (e) $\beta_{X,2} = \log(1)$, $\beta_{U,2} = \log(3)$; (f) $\beta_{X,2} = \log(3)$, $\beta_{U,2} = \log(3)$

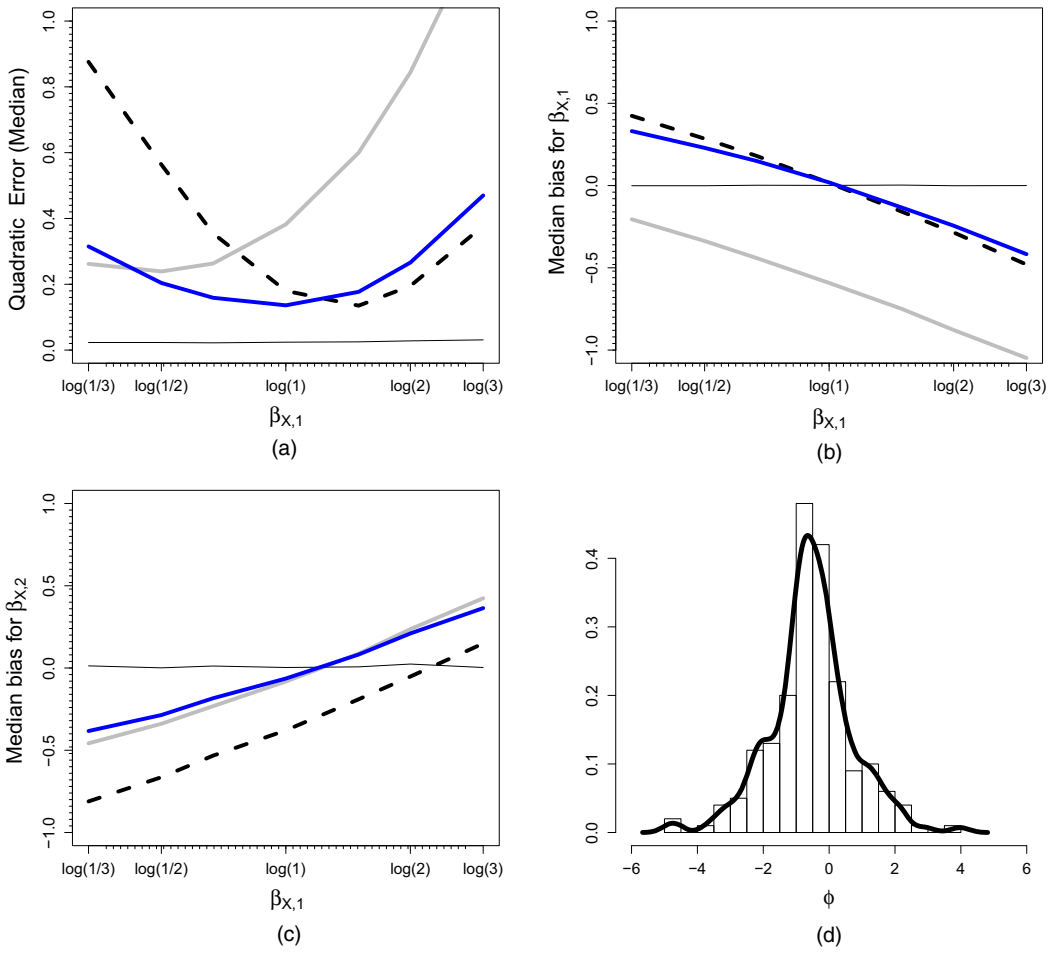


**Fig. 3.** (a) Median of the simulated quadratic errors $E^2 = (\hat{\beta}_{X,1} - \beta_{X,1})^2 + (\hat{\beta}_{X,2} - \beta_{X,2})^2$ for both the survival and the assignment models described in the text ($\beta_{X,2} = \log(3)$; $\beta_{U,2} = \log(3)$), (b) median of the first-period bias ($\hat{\beta}_{X,1} - \beta_{X,1}$) ($\beta_{X,2} = \log(3)$; $\beta_{U,2} = \log(3)$), (c) median of the second-period bias ($\hat{\beta}_{X,2} - \beta_{X,2}$), ($\beta_{X,2} = \log(3)$; $\beta_{U,2} = \log(3)$) (———, real; ———, naive; — —, 2SRI; ———, 2SRI-F) (computations are based on 2000 Monte Carlo iterations for the four models; $\beta_{u,1} = \frac{1}{3}$) and (d) histogram of the variable $\phi$ for one of the simulations (the sample size is 1000)

benefit the most from each procedure. However, the data are exposed to a plethora of selection biases, raising concerns that naive analyses will yield biased results. Because the outcomes of most interest are events such as stroke or death that can occur at any point during follow-up, and such events are often thought to have different risk profiles after the acute period in the first 30 days after the surgery, which is often referred to as the *perioperative period*, these data are ideal for application of 2SRI-F with time-period-specific treatment effects.

We employed 2SRI-F with time-dependent coefficients to estimate the comparative effectiveness of endarterectomy, CEA, *versus* carotid stenting, CAS, the two surgical procedures that are used to intervene on patients with carotid stenosis. The data consist of 73312 patients who received CEA and 12705 who received CAS, between 15 and 89 years of age, over 2003–2016. During the follow-up a total of 8005 events (6600 in CEA) were collected, 730 of them (530 in CEA) during the first 30 days. Fig. 4 shows the short-term Kaplan–Meier estimates for the

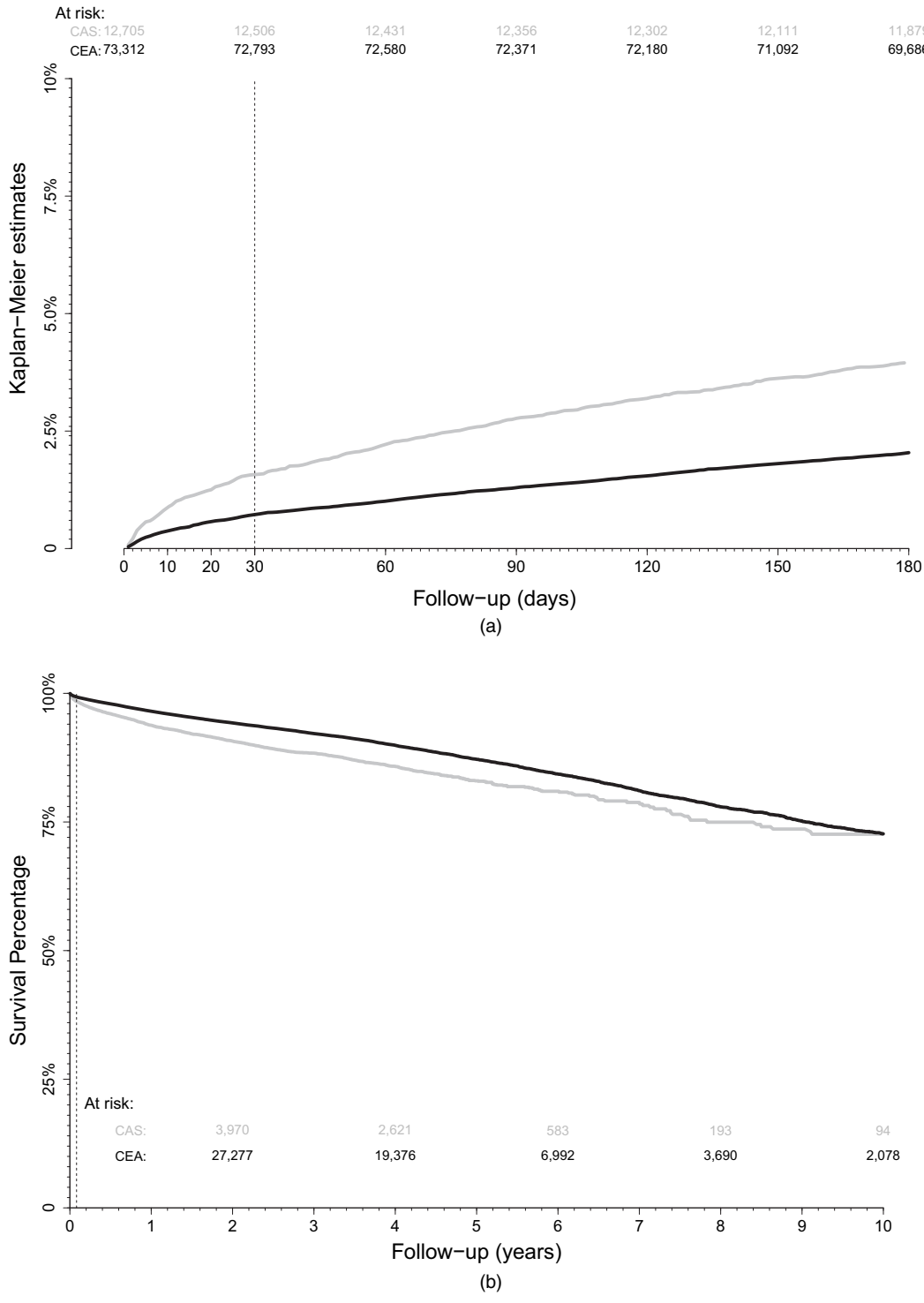


**Fig. 4.** (a) Kaplan–Meier cumulative distribution function and (b) survival percentage estimates by treatment group

cumulative distribution function (Fig. 4(a)) and the long-term Kaplan–Meier survival percentage (Fig. 4(b)) estimates for both the CEA and the CAS groups.

The effect of the procedure received on all-cause mortality risk seems to be stronger in the acute period (hazard ratio (HR) 0.46; 95% confidence interval (0.39; 0.54)) than in the long-term follow-up (HR 0.71 (0.67; 0.76)). When we adjust for observed potential confounders (see Martínez-Camblor *et al*. (2019)) the estimated HRs are only minimally affected: 0.48 (0.40; 0.56) and 0.72 (0.68; 0.77) in the acute and long-term periods respectively. However, the presence of an unmeasured confounder may bias these estimates. We consider as an IV the centre level relative frequency of CEA *versus* CAS procedures over the 12 months before the current patient, i.e. number of CEA/(number of CEA + number of CAS). This IV estimates the propensity of each particular centre to perform CEA. It is justified as an instrument because

(a) hospitals that perform a high relative amount of a certain procedure in the past are likely to keep doing so,
(b) there should be no effect of the relative frequency of CEA *versus* CAS on a patient outcome except through its effect on treatment choice for that patient and
(c) we know of no factors that would influence both the relative frequency of CEA and a patient's outcome.

Reasons (b) and (c) are contingent on adjusting for the total number of CEA and CAS procedures performed at the centre over the previous 12 months, which accounts for the collective experience of the centre at treating carotid artery disease.

On the vascular quality initiative data the IV is highly associated with the choice of treatment, confirming that condition 1 is satisfied. The probability that a randomly selected subject undergoing CEA has a larger value of the instrument than a randomly selected subject undergoing CAS was 0.809 (95% confidence interval (0.805; 0.813)). This IV was uncorrelated with each of the measured confounders, suggesting anecdotally that it may also be uncorrelated with any unmeasured confounder. Hence, it is reasonable to assume that the relationship of the instrument with mortality is solely due to its relationship with the treatment. Fig. 5(a) shows the histogram and Fig. 5(b) the boxplot for the IV in both the CEA and the CAS groups.

When the cut-off point is omitted and we assume that the treatment effect does not change along the follow-up, CEA has some advantage over CAS (HR 0.68 (0.64; 0.71)) which is slightly diluted in the adjustment model (HR 0.69 (0.65; 0.73)). The effect that was reported by 2SRI-F is still significant: 0.81 (0.69; 0.96), but considerably more modest.

The effect of CEA *versus* CAS is more extreme when we allow the treatment effect to change with time. Standard 2SRI (with the total number of procedures per hospital in the previous 12 months as an additional covariate) reports HRs of 0.57 (0.46; 0.72) for the acute period and a non-significant 0.87 (0.74; 1.03) ($p$-value 0.103) for the long-term follow-up. 2SRI-F yields the same effect estimate reported by the unadjusted and adjusted procedures in the acute period (HR 0.48 (0.31; 0.74)) and a non-significant effect in the long-term period (HR 0.89 (0.74; 1.07) with a $p$-value of 0.220). These results were the same under both normal- and gamma-distributed frailties. Table 1 shows the results reported by procedure.

The naive results, ignoring the potential effect of unmeasured confounders, imply a clear change in the survival risk during and after the acute period. Although CEA has a clear advantage in the first 30 days after the surgery (HR 0.46; 95% confidence interval (0.39; 0.54)) the benefit decreases over long-term follow-up (HR 0.71 (0.67; 0.76)). The standard 2SRI procedure reduces the effect in the first period whereas the frailty correction preserves it. None of the IV procedures find significant effects in the long-term period. Note that our Monte Carlo

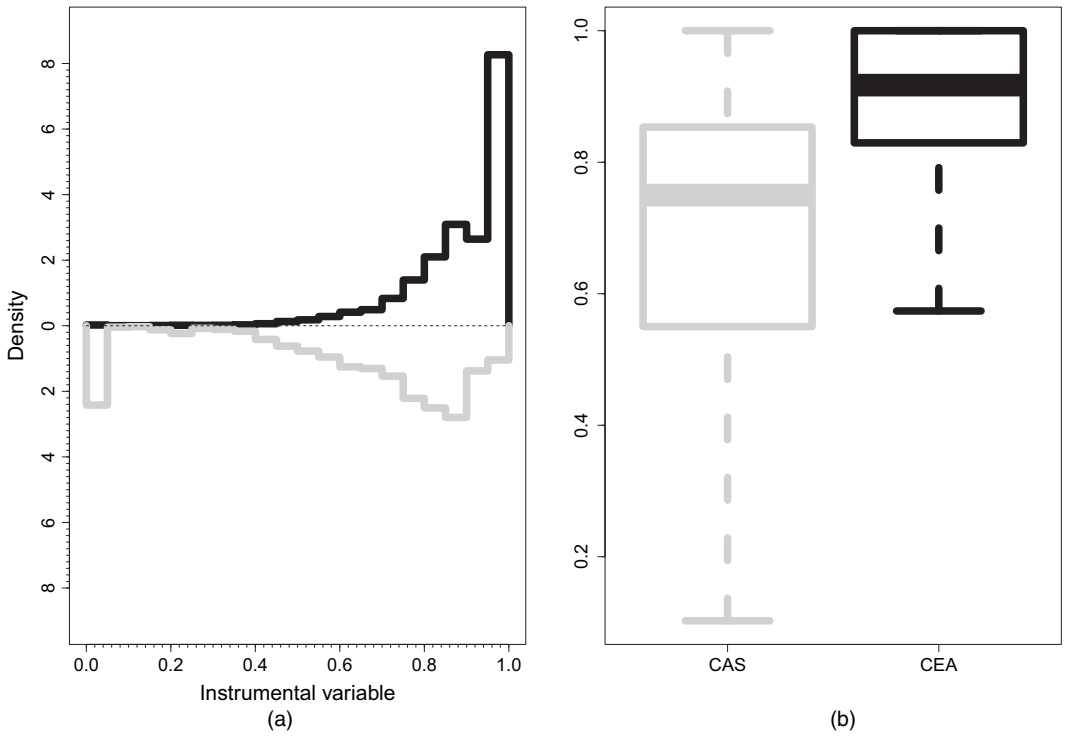


**Fig. 5.** (a) Histograms (———, CEA; ———, CAS) and (b) boxplot for the distribution of IV by treatment group

**Table 1.** HRs and 95% confidence intervals by estimation method

| *Method* | *HR (95% confidence interval)* | | |
|---|---|---|---|
| | *Acute period* | *Long-term follow-up* | *Proportional hazard assumption* |
| Unadjusted Cox model | 0.46 (0.39; 0.54) | 0.71 (0.67; 0.76) | 0.68 (0.64; 0.71) |
| Adjusted Cox model (naive) | 0.48 (0.40; 0.56) | 0.72 (0.68; 0.77) | 0.69 (0.65; 0.73) |
| 2SRI | 0.57 (0.46; 0.72) | 0.87 (0.74; 1.03) | 0.83 (0.70; 0.97) |
| 2SRI-F (Gaussian) | 0.48 (0.31; 0.74) | 0.89 (0.75; 1.07) | 0.81 (0.69; 0.96) |
| 2SRI-F (gamma) | 0.48 (0.31; 0.74) | 0.89 (0.75; 1.07) | 0.79 (0.65; 0.93) |

†Proportional hazard assumption stands for the HRs fixed along the follow-up.

simulations suggest that both 2SRI and 2SRI-F obtained similar results when there is no effect of the treatment, but 2SRI-F improves on the results of 2SRI when the effect is strong. In this case the threshold for a strong treatment effect is close to $\log(\frac{1}{2})$ (see Fig. 2).

## 6. Discussion

Several approaches for IV estimation of Cox proportional hazard regression models have been studied recently (see, for instance, Burne and Abrahamowicz (2019), Li *et al.* (2015), MacKenzie

*et al.* (2014), Tchetgen Tchetgen *et al.* (2015), Martinussen *et al.* (2019) or Wang *et al.* (2018)). Martínez-Camblor *et al.* (2019) proved that, in the two-stage residual inclusion procedure, the first stage introduces a frailty in the second stage which, if not considered, leads to biased estimates. Martinez-Camblor *et al.* (2019) proposed to consider in the second stage a frailty Cox regression model and, via Monte Carlo simulations, showed that the new algorithm, 2SRI-F, improves the 2SRI results. This paper extends the methodology to the case when the effect studied is time dependent. In the simplest case, in which there is a known point of change, there exist different possibilities: to include one or two instruments and to include an independent, shared or correlated univariate frailty. In this paper, we deal with the estimation of a time-dependent coefficient (time modified) in the presence of unmeasured confounding. The case in which there is greater than one time when the treatment effect changes is straightforward. When the time points are not known, a useful strategy is to explore the estimated unadjusted survival curves and/or to make a discrete grid of potential change points and, joint with clinicians' insight, to select those producing relevant changes in the HRs.

Theoretical developments show that the Cox model on which model parameters should be estimated require only a single IV but the residual from the estimated treatment selection equation and a univariate frailty both with time-dependent coefficients must be added to the second-stage equation. We develop an easy-to-implement approximate estimation procedure to overcome the lack of available software for obtaining exact estimates. Our procedure estimates the individual frailty in the first period and adds it as a covariate in the second period, satisfying the constraint that the frailty does not itself vary with time but only its effect varies. To deal with the estimation errors, an additional independent frailty is added in the second period (see Appendix A for an illustrative implementation). The key finding is that this approach works better than two alternative extensions for 2SRI-F that practitioners might first think of as appropriate for the time varying treatment effect case (see Fig. 8 in Appendix A.1). In future work, customized software should be developed for this particular problem to enable the exact inference procedure and potentially improvements could be made to our easy-to-compute estimation procedure.

Monte Carlo simulations reveal that the 2SRI-F procedure performs well in this context. For continuous treatments, it improves the median quadratic error $E^2$, obtained for the 2SRI method in all the situations considered, especially, in the most extreme cases (strongest changes in the coefficients). The results with a binary treatment are also good; however, in this case, the theoretical assumptions do not strictly hold. In the scenario considered, condition 3 is violated and some bias is introduced in the second stage. This bias does not depend strictly on the frailty term but on the relationship between the first-stage residuals and the IV. Recall that the second-stage procedure conditions on $X$ and then $U$ is partially controlled. This theoretical problem is always present in two-stage IV procedures with binary treatments. Note that, with a binary treatment, outcomes frequently do not provide enough information to overt problems like those that occur with weak instruments. Even when we know the correct model from which the outcome was generated and the value of the covariates, we cannot reproduce the value of a single unknown covariate: just a predictive value. Further research about how to deal with this problem is required.

Another interesting point is how to transform the results to survival differences. The new procedure produces good *treatment effect* estimates but, without additional assumptions, it fails in the estimation of the rest of the parameters involved in the model. In addition, as highlighted in Fig. 6 in Appendix A.1, survival curve differences behave differently from HR estimates. Because of the relevance of this issue for effective communication of treatment risk to patients, the transformation of HRs to more meaningful scales should also be the focus of future research.

In the CEA *versus* CAS comparative analysis, naive results imply a clear change in the survival risk during and after the acute period. Although endarterectomy has a clear advantage in the first 30 days after the surgery (HR 0.46; 95% confidence interval (0.39; 0.54)) the benefit decreases over long-term follow-up (HR 0.71 (0.67; 0.76)). The standard 2SRI procedure reduces the effect in the first period whereas the frailty correction preserves it. None of the IV procedures find significant effects in the long-term period. Note that our Monte Carlo simulations suggest that both 2SRI and 2SRI-F obtained similar results when there is no effect of the treatment, but 2SRI-F improves on the results of 2SRI when the effect is strong. In this case the threshold for a strong treatment effect is close to $\log(\frac{1}{2})$ (see Fig. 2). To ignore the presence of a time varying treatment effect and a change point (i.e. a step function relationship) leads to misinterpretation of the true effect of the treatment, implying that CEA has a small protective effect compared with CAS.

We conclude with the recommendation of using the extension of 2SRI-F to the time-dependent treatment effect context that we developed in this paper, which crucially constrains each individual's frailty to be the same in both periods but allows the coefficients of the first-stage residual and the frailty to change between time periods as well as including an additional frailty to account for estimation error in the second stage. This is a non-obvious extension of the procedure that we developed previously for the proportional hazards case that, because of its remarkably good performance, is a significant methodological advance with broad applicability. There exist some computational wrinkles to solve for obtaining exact estimates of individual frailties (Fig. 7 in Appendix A.1). Another important theoretical issue is to adapt the stage 1 equation to obtain exact (and not just approximate) procedures in the case of binary treatments.

## Acknowledgements

All statements in this paper, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of the Patient-Centered Outcomes Research Institute, its Board of Governors or Methodology Committee. The authors are sincerely grateful to our Patient-Centered Outcomes Research Institute Patient Engagement and Governance Committee for reading a draft of the manuscript and for their efforts supporting the development of the research proposal and conducting the research itself.

This work was supported by Patient-Centered Outcomes Research Institute award ME-1503-28261.

The authors have no conflicts of interest to report.

## Appendix A

### *A.1. Additional figures*

Fig. 6 illustrates the problem of transforming HRs to survival curve differences in the presence of unmeasured confounding. It assumes the survival model with hazard function $\lambda(t|X,U)=\frac{1}{4}t\exp\{\log(2)X+U\}$ where $X$ takes values 0 or 1 and depicts the differences between the curves $S_0(t,u)=\exp\{-\int_0^t \lambda(s|X=0,U=u)\mathrm{d}s\}$ and $S_1(t,u)=\exp\{-\int_0^t \lambda(s|X=1,U=u)\mathrm{d}s\}$ against $u$ for $t\in[0,10]$. The effect of the treatment on the survival difference strongly depends on time $t$ and on the unmeasured confounder values. This effect is smallest for the most extreme values of the survival probabilities, which occur at the greatest and lowest values of $U$.

Fig. 7 reports simulation results for the problem of estimating the HR in a proportional hazard Cox regression model with individual frailties. We consider the simple situation in which the hazard function is of the form $\lambda(t|X,U)=\frac{1}{2}t\exp(\beta_X X+\beta_U U)$ where $X$ follows a standard normal distribution and $U$ is independently drawn from a standard normal distribution (Fig. 7(a)) or a gamma distribution with parameter 1 (Fig. 7(b)). Fig. 7 shows the boxplot for the $\hat{\beta}_X$-values based on 2000 Monte Carlo simulations with sample size $n=500$. The horizontal broken line represents the real value of $\beta_X$ ($=\log(2)$) for various

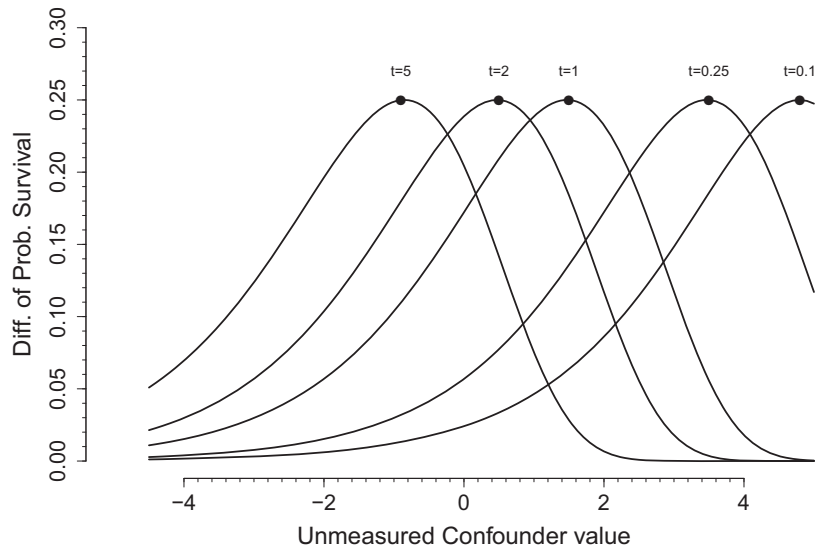


**Fig. 6.** Differences between the curves $S_0(t,u) = \exp\{-\int_0^t \lambda(s|X=0,U=u)\,\mathrm{d}s\}$ and $S_1(t,u) = \exp\{-\int_0^t \lambda(s|X=1,U=u)\mathrm{d}s\}$ against $u$ for a range of $t$-values, where $\lambda(t|X,U) = \frac{1}{4}\,t\exp\{\log(2)X+U\}$ with $X$ taking values 0 or 1

values of $B_U$. The cases that were considered include $X$ and $U$ in the model (real), only include $X$ (naive), only include $X$ with an individual Gaussian frailty and only include $X$ with an individual gamma frailty. Results show that, except for the case $\beta_U = \log(1)$ (no unmeasured covariate), frailty models reduce but do not remove the bias. When the simulations were repeated with a group frailty instead of an individual frailty, we found that the problem reduced but was still present with two members per group but largely disappeared for three or more members per group.

Finally, Fig. 8 reports the results that were obtained by the two alternative extensions of 2SRI introduced in Section 3.2, the algorithms 2SRI-$F_{[1]}$ and 2SRI-$F_{[2]}$, in the first Monte Carlo simulation scenario considered (continuous treatment). In general, 2SRI-F is clearly better than the alternative procedures although they obtain better results when the unmeasured confounding effect is null in the second period.

### *A.2. Computational considerations*

Numerical computation plays an important role in the proportional hazard Cox regression models with univariate frailty (see Fig. 7). There is an increasing number of R packages uploaded to the Comprehensive R Archive Network (`www.r-project.org`) that deal with different approaches to this problem. A non-exhaustive list includes `survival`, `coxme`, `frailtypack`, `frailtyEM`, `frailtyHL`, `frailtySurv`, `parfm` and `PenCoxFrail`. These packages include mainly gamma and Gaussian frailties but other distributions such as the Student $t$-distribution are implemented as well. Parametric survival models are also considered along with univariate, shared, correlated and individual frailties.

The computations in this paper have been implemented by using the `survival` package in R. For simulations, once the parameters have been fixed we draw the standard normal random samples: `eps` ($\epsilon$), $W$ (IV) and $U$ by using `eps=rnorm(n)` (all in the same way). We then compute the treatment value `X=W+U+2eps` for the continuous case and `X=(W+U+eps>0)` for the binary case. Then, given the observed time and the status variables, `(time,exitus)`, we compute the first-stage residuals:

```
res=as.vector(lm(trt~W)$res).
```

We then prepare the data set to be used in the second stage:

```
surv=as.data.frame(cbind(time,exitus,X,res))
surv1=survSplit(Surv(time,exitus)~., surv,cut=t1,episode="timegroup")
```

where `t1` is the change point (equal to 3, in the Monte Carlo simulations).

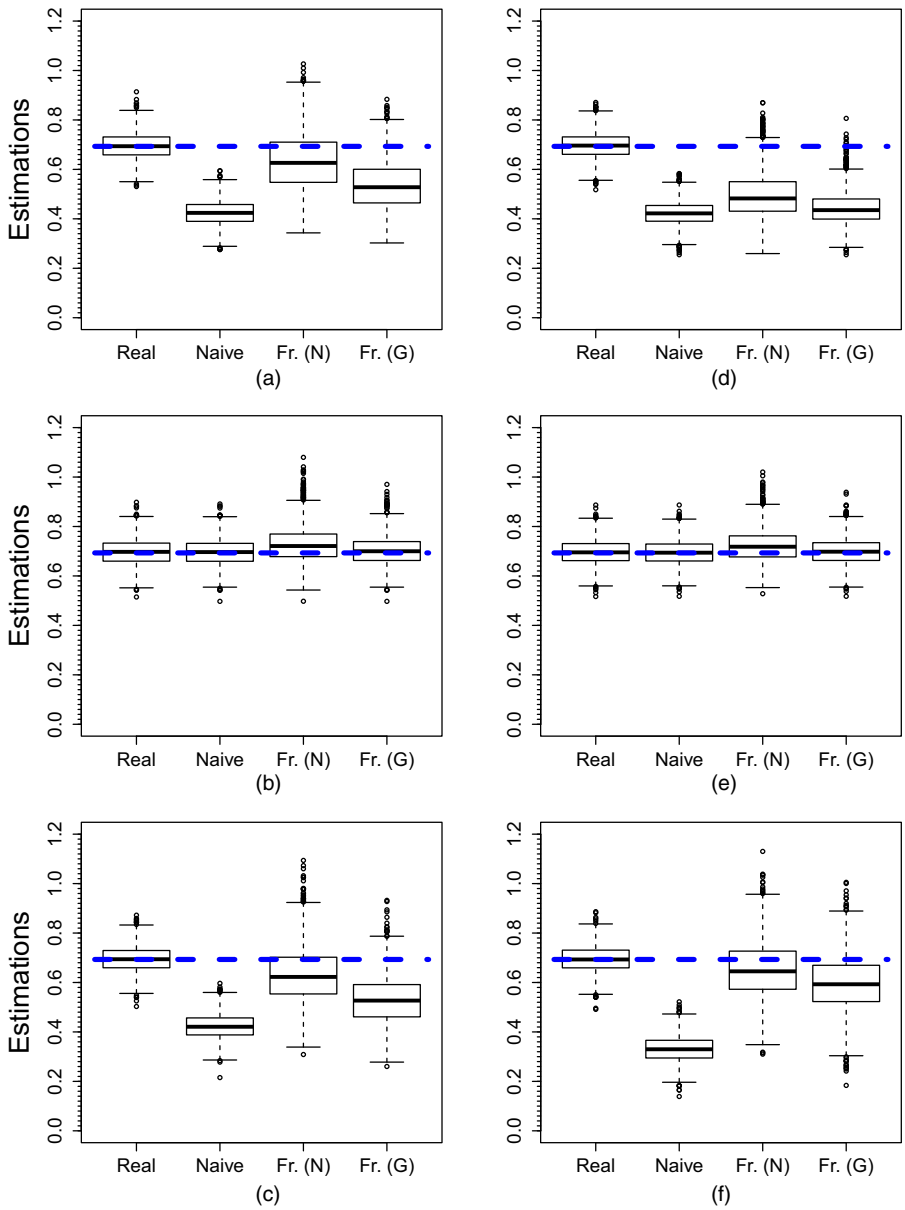


**Fig. 7.** Boxplots for the $\beta_X$-estimations of 2000 Monte Carlo simulations from the hazard function $\lambda(t|X,U) = \frac{1}{2}t\exp(\beta_X X + \beta_U U)$ with $X$ from a standard normal distribution and $U$ independently drawn from (a)–(c) a standard normal distribution or (d)–(f) a gamma distribution with parameter (– – –, true value of $\beta_X (= \log(2))$ (the estimation procedures were to include $X$ and $U$ in the model (real), just to include $X$ (naive), to include $X$ with an individual Gaussian frailty (Fr.(N)) and with an individual gamma frailty (Fr.(G))): (a), (d) $B_U = \log(\frac{1}{3})$; (b), (e) $B_U = \log(1)$; (c), (f) $B_U = \log(3)$

The first period of the second stage is computed from a Cox regression model with Gaussian frailties in which the first-stage residuals are included as covariates:

```
I1<-which(surv1$timegroup==0);n1=length(I1)
M1<-coxph(Surv(start,time,exitus)~trt+res+frailty(1:n1,dist="gaus"),
data=surv1[I1,])
```

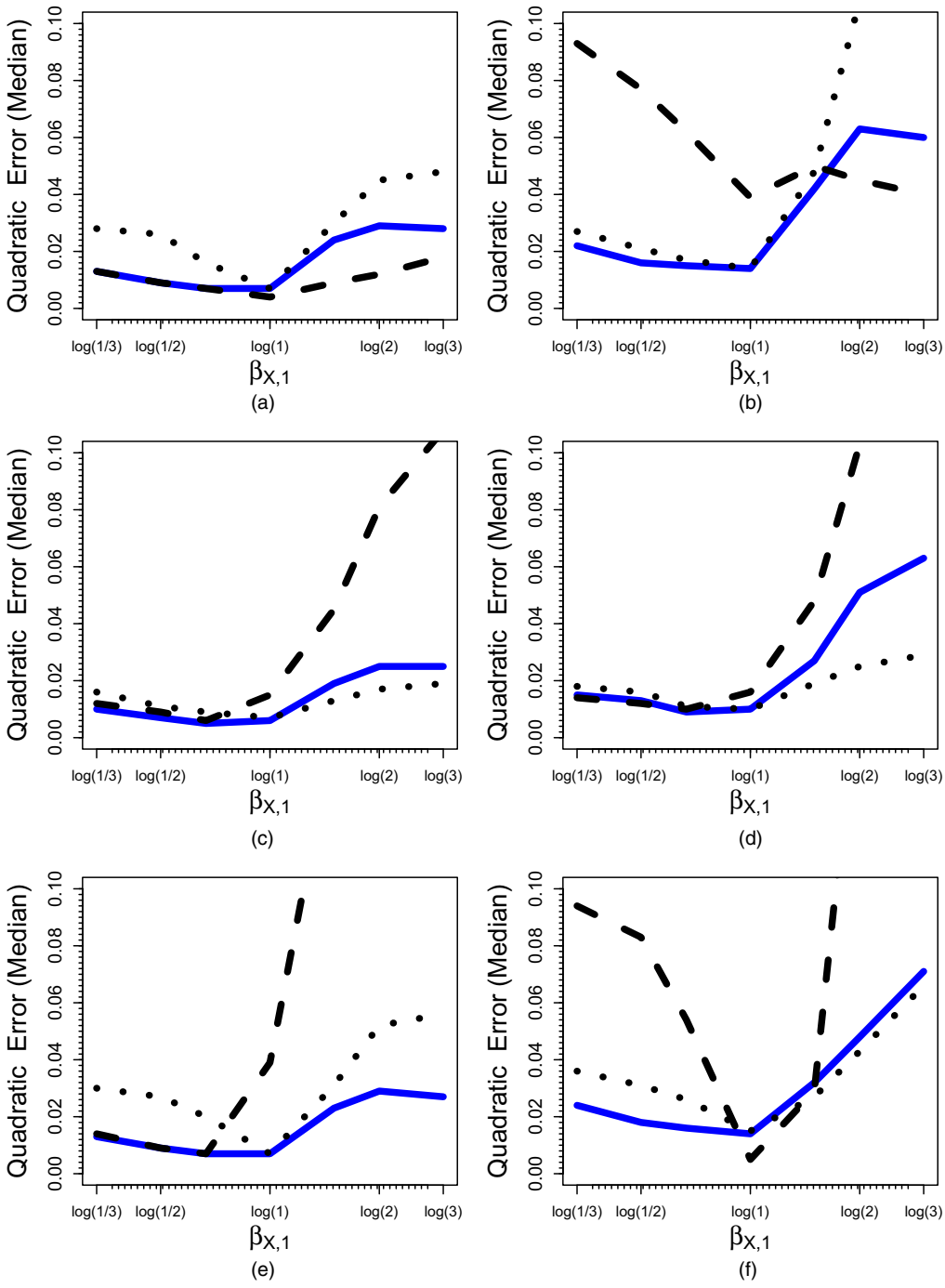


**Fig. 8.** Median of the simulated quadratic errors $E^2 = (\hat{\beta}_{X,1} - \beta_{X,1})^2 + (\hat{\beta}_{X,2} - \beta_{X,2})^2$ for the survival and the 2SRI-F (——), 2SRI-F$_{[1]}$ (· · · · · ·) and 2SRI-F$_{[2]}$ (— —) models (based on 2000 simulated data sets for the three models; the sample size is 1000; $\beta_{U,1} = \frac{1}{3}$): (a) $\beta_{X,2} = \log(1)$, $\beta_{U,2} = \log(\frac{1}{3})$; (b) $\beta_{X,2} = \log(3)$, $\beta_{U,2} = \log(\frac{1}{3})$; (c) $\beta_{X,2} = \log(1)$, $\beta_{U,2} = \log(1)$; (d) $\beta_{X,2} = \log(3)$, $\beta_{U,2} = \log(1)$; (e) $\beta_{X,2} = \log(1)$, $\beta_{U,2} = \log(3)$; (f) $\beta_{X,2} = \log(3)$, $\beta_{U,2} = \log(3)$

We save the individual frailties and the transformed residuals $\hat{R}_1$. Those are included in the second-period model, i.e.

```
F1<-M1$frail
R1<-M1$coef[2]^{-1}*fray+res
surv=as.data.frame(cbind(id,time,exitus,trt,F1,R1))
surv1<-survSplit(Surv(time,exitus)~.,surv,cut=t1,episode="timegroup")
```

The two last lines update the data set to include the new variables. Finally, we compute a Cox regression model with individual Gaussian term including $\hat{R}_1$ as an additional covariate:

```
I2<-which(surv1$timegroup==1);n2=length(I2)
M2<-coxph(Surv(start,time,exitus)~trt+R1+F1+frailty(1:n2,dist="gaus"),
   data=surv1[I2,])
```

## References


Aalen, O. O., Cook, R. J. and Røysland, K. (2015a) Does Cox analysis of a randomized survival study yield a causal treatment effect? *Liftim. Data Anal.*, **21**, 579–593.

Aalen, O. O., Valberg, M., Grotmol, T. and Tretli, S. (2015b) Understanding variation in disease risk: the elusive concept of frailty. *Int. J. Epidem.*, **44**, 1408–1421.

Barker, P. and Henderson, R. (2005) Small sample bias in the gamma frailty model for univariate survival. *Liftim. Data Anal.*, **11**, 265–284.

Burne, R. and Abrahamowicz, M. (2019) Adjustment for time-dependent unmeasured confounders in marginal structural Cox models using validation sample data. *Statist. Meth. Med. Res.*, **28**, 357–371.

Cox, D. R. (1972) Regression models and life-tables. (with discussion). *J. R. Statist. Soc.* B, **34**, 187–220.

Goodney, P. P., Travis, L., Lucas, F. L., Fillinger, M. F., Goodman, D. C., Cronenwett, J. L. and Stone, D. H. (2011) Survival after open versus endovascular thoracic aortic aneurysm repair in an observational study of the medicare population clinical perspective. *Circulation*, **124**, 2661–2669.

Gran, J. M., Hoff, R., Røysland, K., Ledergerber, B., Young, J. and Aalen, O. O. (2018) Estimating the treatment effect on the treated under time-dependent confounding in an application to the Swiss HIV Cohort Study. *Appl. Statist.*, **67**, 103–125.

Greene, W. and Zhang, G. (2003) *Econometric Analysis*. Englewood Cliffs: Prentice Hall.

Hernán, M. (2010) The hazards of hazard ratios. *Epidemiology*, **21**, 13–15.

Hernán, M. and Robins, J. (2006) Instruments for causal inference: an epidemiologist's dream? *Epidemiology*, **17**, 360–372.

Hougaard, P. (1995) Frailty models for survival data. *Liftim. Data Anal.*, **1**, 255–273.

Kosorok, M., Lee, B. L. and Fine, J. (2004) Robust inference for univariate proportional hazards frailty regression models. *Ann. Statist.*, **32**, 1448–1491.

Li, J., Fine, J. and Brookhart, A. (2015) Instrumental variable additive hazards models. *Biometrics*, **71**, 122–130.

MacKenzie, T. A., Tosteson, T. D., Morden, N. E., Stukel, T. A. and O'Malley, A. J. (2014) Using instrumental variables to estimate a Cox's proportional hazards regression subject to additive confounding. *Hlth. Serv. Outoms Res. Methodol.*, **14**, 54–68.

Martínez-Camblor, P., Mackenzie, T., Staiger, D. O., Goodney, P. P. and O'Malley, A. J. (2019) Adjusting for bias introduced by instrumental variable estimation in the Cox proportional hazards model. *Biostatistics*, **20**, 80–96.

Martinussen, T., Nørbo Sørensen, D. and Vansteelandt, S. (2019) Instrumental variables estimation under a structural Cox model. *Biostatistics*, **20**, 65–79.

Martinussen, T. and Vansteelandt, S. (2013) On collapsibility and confounding bias in Cox and Aalen regression models. *Liftim. Data Anal.*, **19**, 279–296.

Murphy, S. A. (1995) Asymptotic theory for the frailty model. *Ann. Statist.*, **23**, 182–198.

Pearl, J. (1995) Causal diagrams for empirical research. *Biometrika*, **82**, 669–688.

Platt, R. W., Schisterman, E. F. and Cole, S. R. (2009) Time-modified confounding. *Am. J. Epidem.*, **170**, 687–694.

Ripatti, S. and Palmgren, J. (2000) Estimation of multivariate frailty models using penalized partial likelihood. *Biometrics*, **56**, 1016–1022.

Robins, J. (1986) A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math. Modllng*, **7**, 1393–1512.

Swanson, S. A. and Hernán, M. A. (2018) The challenging interpretation of instrumental variable estimates under monotonicity. *Int. J. Epidem.*, **47**, 1289–1297.

Tchetgen Tchetgen, E., Walter, S., Vansteelandt, S., Martinussen, T. and Glymour, M. (2015) Instrumental variable estimation in a survival context. *Epidemiology*, **26**, 402–410.

Terza, J. V., Bradford, W. D. and Dismuke, C. E. (2008) The use of linear instrumental variables methods in health services research and health economics: a cautionary note. *Hlth Res. Educ. Trust*, **43**, 1102–1120.

Thanasassoulis, P. and O'Donnell, T. (2009) Mendelian randomization. *J. Am. Med. Ass.*, **301**, 2386–2388.
Therneau, T. M., Grambsch, P. M. and Pankratz, V. S. (2003) Penalized survival models and frailty. *J. Computn/ Graph. Statist.*, **12**, 156–175.
Unkel, S., Farrington, C. P., Whitaker, H. J. and Pebody, R. (2014) Time varying frailty models and the estimation of heterogeneities in transmission of infectious diseases. *Appl. Statist.*, **63**, 141–158.
Wan, F., Small, D., Bekelman, J. E. and Mitra, N. (2015) Bias in estimating the causal hazard ratio when using two-stage instrumental variable methods. *Statist. Med.*, **34**, 2235–2265.
Wang, L., Tchetgen Tchetgen, E., Martinussen, T. and Vansteelandt, S. (2018) Learning causal hazard ratio with endogeneity. *Preprint arXiv:1807.05313*. Harvard University, Cambridge.
Wienke, A. (2010) *Frailty Models in Survival Analysis*. Boca Raton: Chapman and Hall–CRC.