



A goodness-of-fit test for the stratified proportional hazards model for survival data

Rim Ben Elouefi

► To cite this version:

Rim Ben Elouefi. A goodness-of-fit test for the stratified proportional hazards model for survival data. Communications in Statistics - Theory and Methods, 2019, 48 (17), pp.4206-4220. 10.1080/03610926.2018.1489965 . hal-01167977

HAL Id: hal-01167977

<https://hal.science/hal-01167977>

Submitted on 25 Jun 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A goodness-of-fit test for the stratified proportional hazards model for survival data

Rim BEN ELOUEFI

Computational Mathematics Lab, Monastir University, Tunisia and IRMAR-INSa, Rennes, France. Email: Rim.Ben-Elouefi@insa-rennes.fr

Jean-François DUPUY†

IRMAR-INSa, Rennes, France. Email: Jean-Francois.Dupuy@insa-rennes.fr

Abstract. Goodness-of-fit testing is addressed in the stratified proportional hazards model for survival data. A test statistic based on within-strata cumulative sums of martingale residuals over covariates is proposed and its asymptotic distribution is derived under the null hypothesis of model adequacy. A Monte Carlo procedure is proposed to approximate the critical value of the test. Simulation studies are conducted to examine finite-sample performance of the proposed statistic.

Keywords: asymptotic distribution, martingale, residuals, simulations, survival analysis

1. Introduction

The stratified proportional hazards model (see for example [2, 17]) generalizes the usual Cox proportional hazards regression model (see [4]) for survival data by allowing different groups (called strata) of the population under study to have distinct baseline hazard functions. Precisely, in the stratified model, the strata divide the sample individuals into J disjoint groups, each having a distinct baseline hazard function $\alpha_{0,j}$ but a common value for the regression parameter. The hazard function of an individual in stratum j thus takes the form

$$\alpha_j(t) = \alpha_{0,j}(t)e^{\beta_0^\top \mathbf{Z}}, \quad j = 1, \dots, J, \quad (1)$$

where \mathbf{Z} is a p -vector of covariates, β_0 is a p -vector of unknown regression parameters of interest, \top denotes the transpose and $\{\alpha_{0,j}(t) : t \geq 0, j = 1, \dots, J\}$ are J unknown baseline hazard functions. A consistent and asymptotically normal estimator of β_0 is obtained by maximizing the so-called partial likelihood function (see [5]). The partial likelihood for the stratified model (1) is the product over strata of the within-stratum partial likelihoods (see [2]). Further discussion

†Corresponding author: Jean-Francois.Dupuy@insa-rennes.fr

of this model, including estimation of the cumulative baseline hazard functions $A_{0,j}(t) = \int_0^t \alpha_{0,j}(u) du$, can be found in [1] and [17].

The stratified proportional hazards model recently began to be increasingly used in various fields, such as economy, marketing, medicine and public health. For example, in [14], authors use this model to assess the relative performance of piggyback loans and insured loans with respect to residential mortgage life-times in USA. In [11], the author uses a stratified proportional hazards model to identify risk factors of churn among customers of a mobile phone operator. In [22], a stratified proportional hazards model is used to evaluate mortality after radical prostatectomy. In [20], association between exposure to antibiotics in fetal and early life and asthma in childhood is also investigated by using a stratified model.

Despite this recent interest, methodological developments for the stratified proportional hazards model are still a few. In [6] and [10], authors investigate estimation in the stratified model with covariate missing values and covariate measurement error respectively. The prognostic ability of the stratified model is assessed in [3] and [19]. In [13], authors propose a measure to quantify the partial dependence between a survival time and a covariate in the stratified model. Confidence intervals for the difference of median survival times in the stratified model are investigated in [15]. Finally, a stratified proportional hazards model with spatio-temporal heterogeneity is developed in [12]. In the present paper, we consider goodness-of-fit testing for the stratified proportional hazards model.

Many goodness-of-fit tests have been proposed for the usual (unstratified) proportional hazards model. In particular, several authors discussed goodness-of-fit tests based on weighted sums of martingale residuals. A non-exhaustive list of references includes [9, 16, 18, 25], see also [2] for a detailed account of this topic. However, to the best of our knowledge, no goodness-of-fit tests have been proposed yet for the stratified proportional hazards model (1). Our paper aims at filling this gap.

We propose a goodness-of-fit test statistic for model (1) based on within-stratum cumulative sums of residuals. We establish rigorously the asymptotic distribution of this statistic under the null hypothesis that model (1) is correct. A Monte Carlo procedure is proposed to approximate the critical value of the test for a given asymptotic level. Finite-sample performance of the test are investigated via simulations.

The rest of the paper is organized as follows. In Section 2, we construct our test statistic, we investigate its null asymptotic distribution and we discuss Monte Carlo approximation of the critical value. In Section 3, we conduct a simulation study to assess level and power properties of our test for various numbers of strata, sample sizes and proportions of censored observations. Section 4 concludes the paper with some perspectives. Proofs of some intermediate technical results are given in an appendix.

2. The proposed test statistic and decision rule

2.1. Preliminaries and notations

All random variables are defined on the same probability space $(\Omega, \mathcal{C}, \mathbb{P})$. Let T be a random failure time whose distribution depends on a vector of covariates $\mathbf{Z} = (Z_1, \dots, Z_p)^\top \in \mathbb{R}^p$ and on a stratum indicator $S \in \{1, \dots, J\}$. We assume that conditionally on \mathbf{Z} and $S = j$, the hazard function of T is given by model (1). We suppose that T is randomly right-censored by a positive random variable C and that T and C are independent conditionally on \mathbf{Z} and S . The analysis is restricted to the time interval $\mathcal{T} := [0, \tau]$, where $\tau < \infty$ denotes the end of the study. Therefore, we actually observe the duration $X = \min(T, \min(C, \tau))$ and a censoring indicator $\Delta = 1_{\{T \leq \min(C, \tau)\}}$.

Available data consist of n independent and identically distributed copies $(X_i, \Delta_i, \mathbf{Z}_i, S_i)$ of the random vector $(X, \Delta, \mathbf{Z}, S)$. For every $i = 1, \dots, n$ and $t \in \mathcal{T}$, we denote by $N_i(t) = 1_{\{X_i \leq t\}} \Delta_i$ and $Y_i(t) = 1_{\{X_i \geq t\}}$ the failure counting and at-risk processes of the i -th individual respectively. The process $N_i(t)$ has intensity $Y_i(t) \sum_{j=1}^J \alpha_{0,j}(t) e^{\beta_0^\top \mathbf{Z}_i} 1_{\{S_i=j\}}$ with respect to the filtration $(\mathcal{F}_{t,i})_{t \geq 0}$ defined by $\mathcal{F}_{t,i} = \sigma\{\mathbf{Z}_i, S_i, N_i(s), Y_i(s): 0 \leq s \leq t\}$, so that processes

$$M_i(t) = N_i(t) - \int_0^t Y_i(s) \sum_{j=1}^J \alpha_{0,j}(s) e^{\beta_0^\top \mathbf{Z}_i} 1_{\{S_i=j\}} ds, i = 1, \dots, n$$

are martingales.

As mentioned above, a consistent and asymptotically normal estimator $\hat{\beta}_n$ of β_0 in model (1) is obtained by maximizing the partial likelihood

$$\prod_{j=1}^J \prod_{i=1}^n \left(\frac{e^{\beta^\top \mathbf{Z}_i}}{\sum_{k=1}^n Y_k(X_i) e^{\beta^\top \mathbf{Z}_k} 1_{\{S_k=j\}}} \right)^{\Delta_i 1_{\{S_i=j\}}},$$

(see [2]) while cumulative baseline hazard functions $A_{0,j}(t) = \int_0^t \alpha_{0,j}(u) du$ can be estimated by Breslow-type estimators

$$\hat{A}_{0,j}(t) = \int_0^t \frac{\sum_{i=1}^n 1_{\{S_i=j\}} dN_i(s)}{\sum_{k=1}^n Y_k(s) e^{\hat{\beta}_n^\top \mathbf{Z}_k} 1_{\{S_k=j\}}}, \quad j = 1, \dots, J.$$

Martingale residuals in model (1) can be defined similarly as in the classical unstratified Cox model. Precisely, we define the residual martingale for the i -th individual as

$$\widehat{M}_i(t, \hat{\beta}_n) = N_i(t) - \sum_{j=1}^J \int_0^t \frac{Y_i(s) e^{\hat{\beta}_n^\top \mathbf{Z}_i} 1_{\{S_i=j\}}}{S_j^{(0)}(s, \hat{\beta}_n)} d\bar{N}_j(s), \quad (2)$$

where $\bar{N}_j(s) = \sum_{i=1}^n N_i(s) 1_{\{S_i=j\}}$ and $S_j^{(0)}(s, \hat{\beta}_n) = \sum_{i=1}^n Y_i(s) e^{\hat{\beta}_n^\top \mathbf{Z}_i} 1_{\{S_i=j\}}$.

2.2. The proposed test statistic and its asymptotic distribution under H_0

We construct a test statistic for the null hypothesis that the stratified proportional hazards model (1) is true, that is, for

$$H_0 : \alpha_j(t) = \alpha_{0,j}(t)e^{\beta_0^\top \mathbf{Z}}, \quad \text{for } j = 1, \dots, J.$$

Our test statistic is based on within-stratum partial cumulative sums of the martingale residuals (2). So far, model-checking techniques based on cumulative sums of residuals have been developed for various models such as the generalized linear model (*e.g.*, [23, 24]), Cox's model (*e.g.*, [2, 9, 16, 18, 25]) and Aalen's additive risk model ([8]). For the stratified proportional hazards model (1), we propose to consider the process

$$Q_{n,\mathbf{z}}^j(t, \hat{\beta}_n) := \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{M}_i(t, \hat{\beta}_n) 1_{\{S_i=j\}} 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}}, \quad (3)$$

where $\mathbf{z} = (z_1, \dots, z_p)^\top \in \mathbb{R}^p$ and the event $\{\mathbf{Z}_i \leq \mathbf{z}\}$ means that all components $Z_{1,i}, \dots, Z_{p,i}$ of \mathbf{Z}_i are less than or equal to the corresponding components of \mathbf{z} . Under model (1), we have

$$\sum_{i=1}^n \widehat{M}_i(t, \hat{\beta}_n) 1_{\{S_i=j\}} = 0$$

for every $j = 1, \dots, J$. Under H_0 , one would thus expect that the partial-sum process $Q_{n,\mathbf{z}}^j(t, \hat{\beta}_n)$ fluctuates around 0 and a large value of $Q_{n,\mathbf{z}}^j(t, \hat{\beta}_n)$ should lead to the conclusion that model (1) is misspecified for the j -th stratum. Note that the function $\mathbf{z} \mapsto Q_{n,\mathbf{z}}^j(t, \hat{\beta}_n)$ has possible jumps at the distinct values of the \mathbf{Z}_i , $i = 1, \dots, n$. Therefore it is sufficient to consider a finite number Q of values for the vector \mathbf{z} (this issue is discussed in Section 3). If $\mathbf{z}_1, \dots, \mathbf{z}_Q$ are Q such values, we consider the within-stratum statistics $\max_{q \in \mathcal{Q}} |Q_{n,\mathbf{z}_q}^j(t, \hat{\beta}_n)|$ (for $j = 1, \dots, J$), where $\mathcal{Q} = \{1, \dots, Q\}$ and we combine all within-stratum statistics in the overall test statistic

$$\max_{1 \leq j \leq J} \max_{q \in \mathcal{Q}} |Q_{n,\mathbf{z}_q}^j(t, \hat{\beta}_n)|.$$

We consider this statistic at $t = \tau$ in order to use the whole data information and finally, we define our goodness-of-fit test statistic as:

$$\mathcal{S}_n := \max_{1 \leq j \leq J} \max_{q \in \mathcal{Q}} |Q_{n,\mathbf{z}_q}^j(\tau, \hat{\beta}_n)|. \quad (4)$$

We now establish the asymptotic distribution of \mathcal{S}_n under the null hypothesis H_0 . We first introduce some further notations and some regularity conditions. If u is a vector in \mathbb{R}^p , let $u^{\otimes 0} = 1$ and $u^{\otimes 1} = u$. For $k = 0, 1$, $j = 1, \dots, J$, $s \in \mathcal{T}$

and $\beta \in \mathbb{R}^p$, we define

$$\begin{aligned} S_j^{(k)}(s, \beta) &= \sum_{i=1}^n Y_i(s) \mathbf{Z}_i^{\otimes k} e^{\beta^\top \mathbf{Z}_i} 1_{\{S_i=j\}}, \\ S_{j,\mathbf{z}}^{(k)}(s, \beta) &= \sum_{i=1}^n Y_i(s) \mathbf{Z}_i^{\otimes k} e^{\beta^\top \mathbf{Z}_i} 1_{\{S_i=j\}} 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}}, \\ s_j^{(k)}(s, \beta) &= \mathbb{E} \left(Y_i(s) \mathbf{Z}_i^{\otimes k} e^{\beta^\top \mathbf{Z}_i} 1_{\{S_i=j\}} \right), \\ s_{j,\mathbf{z}}^{(k)}(s, \beta) &= \mathbb{E} \left(Y_i(s) \mathbf{Z}_i^{\otimes k} e^{\beta^\top \mathbf{Z}_i} 1_{\{S_i=j\}} 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}} \right). \end{aligned}$$

Finally, $\|\cdot\|_{\mathcal{T}}$ and $\|\cdot\|_{\mathcal{T} \times \mathcal{B}}$ will denote uniform norms on \mathcal{T} and $\mathcal{T} \times \mathcal{B}$ respectively. The asymptotic distribution of our goodness-of-fit statistic \mathcal{S}_n will be established under the following conditions:

C1 β_0 belongs to the interior of a compact subset \mathcal{B} of \mathbb{R}^p .

C2 The time τ is such that $\int_0^\tau \alpha_{0,j}(u) du < \infty$ for all $j = 1, \dots, J$.

C3 The covariate \mathbf{Z} is bounded.

C4 For $j = 1, \dots, J$ and every $\mathbf{z} \in \mathbb{R}^p$, the families $\{s_{j,\mathbf{z}}^{(k)}(s, \cdot); s \in \mathcal{T}\}$ and $\{s_j^{(k)}(s, \cdot); s \in \mathcal{T}\}$ ($k = 0, 1$) of functions of β are equicontinuous at β_0 .

C5 For $j = 1, \dots, J$, there exists a constant $c > 0$ such that $\inf_{s \in \mathcal{T}} s_j^{(0)}(s, \beta_0) > c$.

Our result is as follows:

THEOREM 1. *Let $Q \in \mathbb{N}$ and $\mathbf{z}_1, \dots, \mathbf{z}_Q$ be Q vectors in \mathbb{R}^p . Let \mathbb{W} be a $(J \cdot Q)$ -dimensional Gaussian vector with components denoted by $\mathbb{W} := (\mathbb{W}^{1^\top}, \dots, \mathbb{W}^{J^\top})^\top$, where $\mathbb{W}^j := (W_1^j, \dots, W_Q^j)^\top$ for $j = 1, \dots, J$. Assume that \mathbb{W} has mean zero and variance-covariance matrix \mathbb{S} , where \mathbb{S} is the $(J \cdot Q) \times (J \cdot Q)$ block-matrix*

$$\begin{pmatrix} \Sigma^1 & \dots & 0_{Q \times Q} \\ \vdots & \ddots & \vdots \\ 0_{Q \times Q} & \dots & \Sigma^J \end{pmatrix}$$

and $\Sigma^j = (\Sigma_{k,\ell}^j)_{1 \leq k, \ell \leq Q}$ is the $Q \times Q$ matrix with components

$$\begin{aligned} \Sigma_{k,\ell}^j &:= \mathbb{E} \left[\int_0^\tau \alpha_{0,j}(s) Y_i(s) e^{\beta_0^\top \mathbf{Z}_i} 1_{\{S_i=j\}} 1_{\{\mathbf{Z}_i \leq \mathbf{z}_k\}} 1_{\{\mathbf{Z}_i \leq \mathbf{z}_\ell\}} ds \right] \\ &\quad - \int_0^\tau \alpha_{0,j}(s) \frac{s_{j,\mathbf{z}_k}^{(0)}(s, \beta_0) \cdot s_{j,\mathbf{z}_\ell}^{(0)}(s, \beta_0)}{s_j^{(0)}(s, \beta_0)} ds \end{aligned}$$

and $0_{Q \times Q}$ is a $Q \times Q$ matrix with all components zero. Assume that conditions C1-C5 hold. Then under H_0 , the test statistic \mathcal{S}_n converges in distribution, as n tends to infinity, to $\mathcal{S} := \max_{1 \leq j \leq J} \max_{q \in \mathcal{Q}} |W_q^j|$.

Proof. We need two intermediate lemmas. Their proofs are postponed to an appendix.

LEMMA 1. Let $\mathbf{z}_1, \dots, \mathbf{z}_Q$ be Q vectors in \mathbb{R}^p . Assume that conditions C1-C5 hold and for $j = 1, \dots, J$, let $\mathbb{Q}_n^j(\cdot, \beta) := (Q_{n, \mathbf{z}_1}^j(\cdot, \beta), \dots, Q_{n, \mathbf{z}_Q}^j(\cdot, \beta))^\top$. Then under H_0 , the process $(\mathbb{Q}_n^1(\cdot, \beta_0)^\top, \dots, \mathbb{Q}_n^J(\cdot, \beta_0)^\top)^\top$ converges weakly in $(D(\mathcal{T}))^{J \cdot Q}$ to a zero-mean Gaussian process $(\mathbb{W}^{1\top}, \dots, \mathbb{W}^{J\top})^\top$ (with $\mathbb{W}^j := (W_1^j, \dots, W_Q^j)^\top$) with covariance function

$$\begin{aligned} \text{cov}(W_k^j(t_1), W_\ell^j(t_2)) = & \mathbb{E} \left[\int_0^{\min(t_1, t_2)} \alpha_{0,j}(s) Y_i(s) e^{\beta_0^\top \mathbf{z}_i} 1_{\{S_i=j\}} 1_{\{\mathbf{z}_i \leq \mathbf{z}_k\}} 1_{\{\mathbf{z}_i \leq \mathbf{z}_\ell\}} ds \right] \\ & - \int_0^{\min(t_1, t_2)} \alpha_{0,j}(s) \frac{s_{j, \mathbf{z}_k}^{(0)}(s, \beta_0) \cdot s_{j, \mathbf{z}_\ell}^{(0)}(s, \beta_0)}{s_j^{(0)}(s, \beta_0)} ds \end{aligned}$$

and $\text{cov}(W_k^j(t_1), W_\ell^{j'}(t_2)) = 0$ if $j \neq j'$ (with $1 \leq k, \ell \leq Q$ and $t_1, t_2 \in \mathcal{T}$).

LEMMA 2. Assume that conditions C1-C5 hold. Then under H_0 , $\mathbb{Q}_n^j(\cdot, \hat{\beta}_n) = \mathbb{Q}_n^j(\cdot, \beta_0) + o_{\mathbb{P}}(1)$ where the $o_{\mathbb{P}}(1)$ is uniform on \mathcal{T} .

We now prove Theorem 1. It follows from Lemma 1 and Lemma 2 that under H_0 , the process $(\mathbb{Q}_n^1(\cdot, \hat{\beta}_n)^\top, \dots, \mathbb{Q}_n^J(\cdot, \hat{\beta}_n)^\top)^\top$ converges weakly to a zero-mean Gaussian process on $\mathcal{T}^{J \cdot Q}$, with covariance function as in Lemma 1. Hence $(\mathbb{Q}_n^1(\tau, \hat{\beta}_n)^\top, \dots, \mathbb{Q}_n^J(\tau, \hat{\beta}_n)^\top)^\top$ converges in distribution to a Gaussian vector with mean zero and variance-covariance matrix given by \mathbb{S} in Theorem 1 (this vector is denoted by $\mathbb{W} := (\mathbb{W}^{1\top}, \dots, \mathbb{W}^{J\top})^\top$, with $\mathbb{W}^{j\top} := (W_1^j, \dots, W_Q^j)$, $j = 1, \dots, J$). By the continuous mapping theorem, \mathcal{S}_n converges in distribution to $\max_{1 \leq j \leq J} \max_{q \in Q} |W_q^j|$.

□

Based on Theorem 1, a natural decision rule is as follows: reject H_0 (at the asymptotic level $\alpha \in (0, 1)$) if \mathcal{S}_n is greater than the quantile $q_{1-\alpha}(\mathcal{S})$ of order $1 - \alpha$ of the distribution of \mathcal{S} . However, this quantile - or critical value of the test - is unknown. In the next section, we propose to estimate it via Monte Carlo simulations.

2.3. Monte Carlo estimation of the critical value and decision rule

In order to estimate $q_{1-\alpha}(\mathcal{S})$, we propose to: i) simulate a large number (say M) of realizations of \mathbb{W} , ii) calculate the corresponding values s_1, \dots, s_M of \mathcal{S} , iii) approximate $q_{1-\alpha}(\mathcal{S})$ by the empirical quantile of order $1 - \alpha$ of (s_1, \dots, s_M) . However, the matrix \mathbb{S} needed to simulate data from $\mathbb{W} \sim \mathcal{N}_{(J \cdot Q)}(0, \mathbb{S})$ is unknown. Under H_0 , this matrix can be estimated by replacing unknown terms in

$\Sigma_{k,\ell}^j$ by their empirical counterparts. Precisely, we define:

$$\widehat{\Sigma}_{k,\ell}^j := \int_0^\tau \left\{ \frac{S_{j,\check{\mathbf{z}}_{k,\ell}}^{(0)}(s, \widehat{\beta}_n)}{S_j^{(0)}(s, \widehat{\beta}_n)} - \frac{S_{j,\mathbf{z}_k}^{(0)}(s, \widehat{\beta}_n) \cdot S_{j,\mathbf{z}_\ell}^{(0)}(s, \widehat{\beta}_n)}{(S_j^{(0)}(s, \widehat{\beta}_n))^2} \right\} \frac{d\bar{N}_j(s)}{n},$$

where $\check{\mathbf{z}}_{k,\ell}$ denotes the vector in \mathbb{R}^p whose j -th component ($j = 1, \dots, p$) is the minimum of the j -th components of \mathbf{z}_k and \mathbf{z}_ℓ . This estimator is consistent, as stated in Proposition 2.

PROPOSITION 2. *Assume that conditions C1-C5 hold. Then, under H_0 , $\widehat{\Sigma}_{k,\ell}^j$ converges in probability to $\Sigma_{k,\ell}^j$ as n tends to infinity.*

Proof. Let

$$V_{k,\ell}^j(s, \beta) := \frac{S_{j,\check{\mathbf{z}}_{k,\ell}}^{(0)}(s, \beta)}{S_j^{(0)}(s, \beta)} - \frac{S_{j,\mathbf{z}_k}^{(0)}(s, \beta) \cdot S_{j,\mathbf{z}_\ell}^{(0)}(s, \beta)}{(S_j^{(0)}(s, \beta))^2}$$

and

$$v_{k,\ell}^j(s, \beta) := \frac{s_{j,\check{\mathbf{z}}_{k,\ell}}^{(0)}(s, \beta)}{s_j^{(0)}(s, \beta)} - \frac{s_{j,\mathbf{z}_k}^{(0)}(s, \beta) \cdot s_{j,\mathbf{z}_\ell}^{(0)}(s, \beta)}{(s_j^{(0)}(s, \beta))^2}.$$

We have:

$$\begin{aligned} \left| \widehat{\Sigma}_{k,\ell}^j - \Sigma_{k,\ell}^j \right| &= \left| \int_0^\tau V_{k,\ell}^j(s, \widehat{\beta}_n) \frac{d\bar{N}_j(s)}{n} - \int_0^\tau v_{k,\ell}^j(s, \beta_0) s_j^{(0)}(s, \beta_0) \alpha_{0,j}(s) ds \right| \\ &\leq \left| \int_0^\tau \left\{ V_{k,\ell}^j(s, \widehat{\beta}_n) - v_{k,\ell}^j(s, \widehat{\beta}_n) \right\} \frac{d\bar{N}_j(s)}{n} \right| \\ &\quad + \left| \int_0^\tau \left\{ v_{k,\ell}^j(s, \widehat{\beta}_n) - v_{k,\ell}^j(s, \beta_0) \right\} \frac{d\bar{N}_j(s)}{n} \right| \\ &\quad + \left| \int_0^\tau v_{k,\ell}^j(s, \beta_0) \left\{ \frac{d\bar{N}_j(s)}{n} - \frac{1}{n} S_j^{(0)}(s, \beta_0) \alpha_{0,j}(s) ds \right\} \right| \\ &\quad + \left| \int_0^\tau v_{k,\ell}^j(s, \beta_0) \left\{ \frac{1}{n} S_j^{(0)}(s, \beta_0) - s_j^{(0)}(s, \beta_0) \right\} \alpha_{0,j}(s) ds \right| \quad (5) \end{aligned}$$

Under conditions C1-C5, $\|V_{k,\ell}^j - v_{k,\ell}^j\|_{\mathcal{T} \times \mathcal{B}}$ converges to zero in probability. Moreover, $\bar{N}_j(\tau)/n$ tends to $\mathbb{E}[N(\tau)1_{\{S=j\}}] < \infty$. Therefore, the first term in (5) converges to zero in probability. Condition C4 and the convergence of $\bar{N}_j(\tau)/n$ to $\mathbb{E}[N(\tau)1_{\{S=j\}}] < \infty$ imply that the second term in (5) also converges to zero. Consider the third term in (5). Let $\delta > 0, \eta > 0$. By Lenglart inequality (e.g., [1]), we have

$$\begin{aligned} \mathbb{P} \left(\left| \int_0^\tau v_{k,\ell}^j(s, \beta_0) \frac{1}{n} d\bar{M}_j(s) \right| > \eta \right) \\ \leq \frac{\delta}{\eta^2} + \mathbb{P} \left(\int_0^\tau (v_{k,\ell}^j(s, \beta_0))^2 \frac{1}{n} S_j^{(0)}(s, \beta_0) \alpha_{0,j}(s) ds > n\delta \right). \end{aligned}$$

For any $\delta > 0$, the probability on the right-hand side of this inequality converges to zero as n tends to infinity, and since δ is arbitrary, $\mathbb{P}(|\int_0^\tau v_{k,\ell}^j(s, \beta_0) \frac{1}{n} d\bar{M}_j(s)| > \eta)$ must converge to zero. Thus the third term in (5) converges to zero in probability. Finally,

$$\begin{aligned} & \left| \int_0^\tau v_{k,\ell}^j(s, \beta_0) \left\{ \frac{1}{n} S_j^{(0)}(s, \beta_0) - s_j^{(0)}(s, \beta_0) \right\} \alpha_{0,j}(s) ds \right| \\ & \leq \int_0^\tau \left| v_{k,\ell}^j(s, \beta_0) \right| \alpha_{0,j}(s) ds \times \left\| \frac{1}{n} S_j^{(0)}(\cdot, \beta_0) - s_j^{(0)}(\cdot, \beta_0) \right\|_{\mathcal{T}}. \end{aligned}$$

The first term on the right-hand side of this inequality is bounded under conditions C1-C5 and the second term converges to zero in probability by Glivenko-Cantelli theorem. Thus the fourth term in (5) converges to zero and the proof is complete.

□

In what follows, we denote by $\widehat{\mathbb{S}}$ the estimator of \mathbb{S} obtained by replacing the $\Sigma_{k,\ell}^j$ by their estimators ($j = 1, \dots, J; k, \ell = 1, \dots, Q$).

We now apply a parametric bootstrap procedure: i) we simulate M vectors $\mathbb{W}_1^*, \dots, \mathbb{W}_M^*$ from $\mathbb{W}^* \sim \mathcal{N}_{(J,Q)}(0, \widehat{\mathbb{S}})$, ii) we calculate the corresponding values $s_i^* = \max_{1 \leq j \leq J} \max_{q \in Q} |W_{i,q}^{*j}|$, $i = 1, \dots, M$, iii) we estimate $q_{1-\alpha}(\mathcal{S})$ by the empirical quantile $q_{1-\alpha}^*$ of order $1 - \alpha$ of s_1^*, \dots, s_M^* . Our decision rule is finally: "reject H_0 at the asymptotic level $\alpha \in (0, 1)$ if $\mathcal{S}_n \geq q_{1-\alpha}^*$ ".

In the next section, we investigate this decision rule via simulations. Several issues are discussed, such as the choice of Q and $\mathbf{z}_1, \dots, \mathbf{z}_Q$.

3. Simulation study

In this simulation study, we investigate the finite sample behaviour of our test statistic for various numbers of strata typically encountered in practice ($J = 3, 5$). Precisely, we assess level and power of our test against various alternatives, for several sample sizes and censoring fractions. Simulations are run using the statistical language R (see [21]) under a Linux Ubuntu Server 14.04 LTS 64Bits with two processors Intel E5 2640v2 running at 1600MHZ, with 256Go RAM.

First, we simulate $N = 10^5$ samples of size n of observations from model (1) with $J = 3$. The sample sizes in the 3 strata are denoted by (n_1, n_2, n_3) , with $n = \sum_{j=1}^3 n_j$. We consider several values for (n_1, n_2, n_3) , namely (100, 110, 80), (150, 175, 120) and (200, 225, 190). The baseline hazard function in stratum j ($j = 1, 2, 3$) is $\alpha_{0,j}(t) = \lambda_j \alpha_j t^{\alpha_j - 1}$ with $(\alpha_1, \lambda_1) = (2.1, 1)$, $(\alpha_2, \lambda_2) = (1.2, 0.75)$, $(\alpha_3, \lambda_3) = (1.8, 1.5)$. We consider a two-dimensional covariate $\mathbf{Z} = (Z_1, Z_2)^\top$, where Z_1 and Z_2 are independent and distributed as a Gaussian $\mathcal{N}(0, 1)$ and a uniform $\mathcal{U}(1, 3)$ respectively. We take $\beta_0 = (0.2, 0.7)^\top$. The model used for simulating data is thus $\alpha_j(t) = \lambda_j \alpha_j t^{\alpha_j - 1} e^{0.2Z_1 + 0.7Z_2}$, $j = 1, 2, 3$. Censoring times are simulated from an exponential distribution with parameter $\mu > 0$, where μ

is chosen to yield some pre-specified proportion c of censored observations (we consider $c = 0.1, 0.2, 0.4$).

As mentioned in Section 2.2, the function $\mathbf{z} \mapsto Q_{n,\mathbf{z}}^j(t, \hat{\beta}_n)$ has possible jumps at the distinct values of the \mathbf{Z}_i , $i = 1, \dots, n$. Therefore it is sufficient to consider a finite number Q of values for the vector \mathbf{z} . One may consider all distinct \mathbf{Z}_i but our numerical experiments showed that no major change affects the outcome of the test when Q is smaller than n , provided that Q stays large enough. Thus in this simulation study, we take $Q = 250$ when $(n_1, n_2, n_3) = (100, 110, 80)$, $Q = 400$ when $(n_1, n_2, n_3) = (150, 175, 120)$, $Q = 600$ when $(n_1, n_2, n_3) = (200, 225, 190)$ and we use a regular grid of values \mathbf{z}_q over the range of the \mathbf{Z}_i .

Under the setting described above, the null hypothesis H_0 holds and we investigate level of our test. For each of the N simulated samples, we calculate the test statistic \mathcal{S}_n and approximate the critical value (for an asymptotic level $\alpha = 0.05$) by using the Monte Carlo procedure described in Section 2.3 (with $M = 5000$). Finally, we apply the proposed decision rule. Table 1 provides empirical level of the test for the various configurations of the simulation design parameters (see the row labeled " H_0 ").

Next, we investigate power of our test. For $J = 3$, we consider the following alternatives (respectively denoted by $H_{1,a}$, $H_{1,b}$, $H_{1,c}$ and $H_{1,d}$):

- *non-proportional hazards model*: the hazard function in stratum j is chosen as $\alpha_j(t) = \alpha_j e^{1.2Z_1 + 1.5Z_2 \times t}$, $j = 1, 2, 3$. We take $\alpha_1 = 0.01, \alpha_2 = 0.1, \alpha_3 = 0.25$. Hazard ratios are not constant in time and thus, the model is not a proportional hazards model.
- *covariate threshold effect*: the hazard function in stratum j is chosen as $\alpha_j(t) = \lambda_j \alpha_j t^{\alpha_j - 1} e^{1.7Z_1 1_{\{Z_1 > \xi_j\}} + 0.5Z_2}$, $j = 1, 2, 3$. We take $(\alpha_1, \lambda_1, \xi_1) = (1.5, 1, 0.6)$, $(\alpha_2, \lambda_2, \xi_2) = (0.5, 0.75, 1)$ and $(\alpha_3, \lambda_3, \xi_3) = (1, 1.25, 0.8)$.
- *distinct regression parameters across strata*: the hazard function in stratum j is chosen as $\alpha_j(t) = \lambda_j \alpha_j t^{\alpha_j - 1} e^{\beta_{1,j}Z_1 + \beta_{2,j}Z_2}$, $j = 1, 2, 3$. We take $(\alpha_1, \lambda_1, \beta_{1,1}, \beta_{2,1}) = (2.1, 1, 0.2, 0.7)$, $(\alpha_2, \lambda_2, \beta_{1,2}, \beta_{2,2}) = (1.2, 0.75, 1, 1)$ and $(\alpha_3, \lambda_3, \beta_{1,3}, \beta_{2,3}) = (1.8, 1.5, 0.2, 0.2)$.
- *omitted covariates*: the hazard function in stratum j is taken as $\alpha_j(t) = \lambda_j \alpha_j t^{\alpha_j - 1} e^{Z_1 - 0.7Z_2 + 0.75Z_3}$ ($j = 1, 2, 3$) where the additional covariate Z_3 is distributed as a $\mathcal{N}(1, 0.25)$. We take $(\alpha_1, \lambda_1) = (2.1, 1)$, $(\alpha_2, \lambda_2) = (1.2, 0.75)$ and $(\alpha_3, \lambda_3) = (1.8, 1.5)$.

We simulate $N = 10^5$ samples of observations under each of these alternatives. The simulation design (within-stratum sample sizes, censoring proportion) is the same as for evaluating level. For each sample and each alternative hypothesis, we calculate \mathcal{S}_n (based on the fitted model $\alpha_{0,j}(t) \exp(\beta_1 Z_1 + \beta_2 Z_2)$, $j = 1, 2, 3$), we approximate the critical value (for an asymptotic level $\alpha = 0.05$) by using the Monte Carlo procedure of Section 2.3 (with $M = 5000$) and finally, we apply

the proposed decision rule. Empirical powers of the test under each alternative are reported in Table 1 (see rows labeled $H_{1,a}, H_{1,b}, H_{1,c}, H_{1,d}$).

	(100, 110, 80)			(150, 175, 120)			(200, 225, 190)		
c	0.1	0.2	0.4	0.1	0.2	0.4	0.1	0.2	0.4
H_0	0.0569	0.0679	0.0629	0.0436	0.0612	0.0629	0.0686	0.0505	0.0650
$H_{1,a}$	0.6762	0.5408	0.4187	0.8177	0.6982	0.4683	0.9705	0.8443	0.5790
$H_{1,b}$	0.6287	0.5320	0.4635	0.7577	0.7370	0.7216	0.9149	0.8600	0.8561
$H_{1,c}$	0.7469	0.7239	0.6648	0.9178	0.9130	0.8563	0.9818	0.9779	0.9659
$H_{1,d}$	0.6763	0.6483	0.4889	0.8423	0.8217	0.7897	0.9357	0.9191	0.8931

Table 1. Empirical size and power of the proposed test for various censoring proportions and sample sizes, with $J = 3$. All results are based on 10^5 simulated samples.

Then, we conduct a similar study as above, with $J = 5$. For investigating level, we simulate data from the stratified model $\alpha_{0,j}(t) = \lambda_j \alpha_j t^{\alpha_j - 1} e^{0.2Z_1 + 0.7Z_2}$ ($j = 1, \dots, 5$) with the same (α_i, λ_i) , $i = 1, 2, 3$ as above and $(\alpha_4, \lambda_4) = (1, 1)$, $(\alpha_5, \lambda_5) = (1.2, 0.5)$. We consider following sample sizes: (100, 110, 80, 110, 70), (150, 175, 120, 80, 110) and (200, 225, 190, 150, 120). A similar procedure as for $J = 3$ yields results in Table 2 (see the row labeled " H_0 "). The power of the test is investigated under the same alternatives as above, with two additional strata. For conciseness, we postpone description of parameters values for strata $j = 4, 5$ to an appendix. Empirical powers of the test under each alternative are reported in Table 2 (see rows labeled $H_{1,a}, H_{1,b}, H_{1,c}, H_{1,d}$).

	(100, 110, 80, 110, 70)				(150, 175, 120, 80, 110)			(200, 225, 190, 150, 120)		
c	0.1	0.2	0.4		0.1	0.2	0.4	0.1	0.2	0.4
H_0	0.0638	0.0612	0.0428		0.0401	0.0584	0.0502	0.0524	0.0624	0.0609
$H_{1,a}$	0.7365	0.5967	0.2355		0.9341	0.8517	0.5581	0.9782	0.9263	0.6105
$H_{1,b}$	0.4502	0.4335	0.3837		0.8079	0.7283	0.6823	0.8876	0.8794	0.8627
$H_{1,c}$	0.8167	0.6616	0.5869		0.9195	0.9003	0.8830	0.9785	0.9718	0.9282
$H_{1,d}$	0.5898	0.5233	0.5211		0.8018	0.7937	0.7955	0.9657	0.9541	0.9210

Table 2. Empirical size and power of the proposed test for various censoring proportions and sample sizes, with $J = 5$. All results are based on 10^5 simulated samples.

From these results, the proposed test statistic performs well under a variety of conditions. The empirical level is close to the nominal level even when censoring

is large (40%, say). As expected, power of the test increases when within-strata sample sizes increase and decreases when censoring increases. However, when censoring is low to moderate (20%, say), the test maintains satisfactory power against every alternative provided that within-strata sample sizes are all sufficiently large (greater than 100, say). When censoring is heavy (40%, say), the test still maintains good power if all within-strata sample sizes are sufficiently large, say 200. Overall, our test appears to provide an efficient tool for assessing adequacy of the stratified proportional hazards model under usual conditions of sample size and censoring.

4. Conclusion

In this paper, we propose a goodness-of-fit test statistic for the stratified proportional hazards model and we establish its asymptotic distribution under the null hypothesis that a stratified model holds. Our simulation study suggests that the proposed test performs well under a wide range of conditions (sample size, censoring fraction, alternative hypothesis).

Now, several issues deserve attention. First, in the stratified proportional hazards model, the variable used for stratifying the population under study is often discrete (*e.g.*, gender, disease stage, socio-professional category. . .). Sometimes, it may be relevant to stratify according to a continuous variable W (such as tumour size, level of salary. . .). For example, in the case of two strata, the resulting model may be written as $\alpha_{0,1}(t) \exp(\beta^\top \mathbf{Z}) 1_{\{W \leq w_0\}} + \alpha_{0,2}(t) \exp(\beta^\top \mathbf{Z}) 1_{\{W > w_0\}}$, where w_0 is some unknown threshold. To the best of our knowledge, no procedure was proposed yet to estimate w_0 . Our test statistic may be used for that purpose. For example, one may test goodness of fit of the model $\alpha_{0,1}(t) \exp(\beta^\top \mathbf{Z}) 1_{\{W \leq w\}} + \alpha_{0,2}(t) \exp(\beta^\top \mathbf{Z}) 1_{\{W > w\}}$ for several values of w (using our test statistic \mathcal{S}_n) and retain the value \hat{w} yielding the less significant result. We are currently exploring this original application of our test statistic.

Second, our test is feasible when the number of covariates stays moderate. A large number of covariates will raise some computational issues. For example, calculating (3) over a fine grid of a high-dimensional space will eventually be time-consuming. We are currently exploring some directions to reduce this computational burden.

Appendix A. Proofs of Lemma 1 and Lemma 2.

Proof of Lemma 1 We use the martingale central limit theorem (*e.g.*, Theorem 5.3.5 in [7]) to establish weak convergence of the process $(\mathbb{Q}_n^1(\cdot, \beta_0)^\top, \dots, \mathbb{Q}_n^J(\cdot, \beta_0)^\top)$. First, we provide an alternative expression for $Q_{n,\mathbf{z}}^j(t, \beta_0)$. For $j = 1, \dots, J$, let

$\bar{N}_j(s) = \sum_{i=1}^n 1_{\{S_i=j\}} N_i(s)$. From (3), we have:

$$\begin{aligned}
Q_{n,\mathbf{z}}^j(t, \beta_0) &:= \frac{1}{\sqrt{n}} \sum_{i=1}^n \widehat{M}_i(t, \beta_0) 1_{\{S_i=j\}} 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}}, \\
&= \sum_{i=1}^n \int_0^t \frac{1}{\sqrt{n}} 1_{\{S_i=j\}} 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}} \left\{ dN_i(s) - \sum_{\ell=1}^J \frac{Y_i(s) e^{\beta_0^\top \mathbf{Z}_i} 1_{\{S_i=\ell\}}}{S_\ell^{(0)}(s, \beta_0)} d\bar{N}_\ell(s) \right\} \\
&= \sum_{i=1}^n \int_0^t \frac{1}{\sqrt{n}} 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}} \left\{ 1_{\{S_i=j\}} dN_i(s) - \frac{Y_i(s) e^{\beta_0^\top \mathbf{Z}_i} 1_{\{S_i=j\}}}{S_j^{(0)}(s, \beta_0)} d\bar{N}_j(s) \right\} \\
&= \sum_{i=1}^n \int_0^t \frac{1}{\sqrt{n}} \left\{ 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}} - \frac{S_{j,\mathbf{z}}^{(0)}(s, \beta_0)}{S_j^{(0)}(s, \beta_0)} \right\} 1_{\{S_i=j\}} dN_i(s).
\end{aligned}$$

Now, recall that $N_i(t) = \int_0^t Y_i(s) \sum_{j=1}^J \alpha_{0,j}(s) e^{\beta_0^\top \mathbf{Z}_i} 1_{\{S_i=j\}} ds + M_i(t)$ and thus:

$$\begin{aligned}
Q_{n,\mathbf{z}}^j(t, \beta_0) &= \sum_{i=1}^n \int_0^t \frac{1}{\sqrt{n}} \left\{ 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}} - \frac{S_{j,\mathbf{z}}^{(0)}(s, \beta_0)}{S_j^{(0)}(s, \beta_0)} \right\} 1_{\{S_i=j\}} Y_i(s) \alpha_{0,j}(s) e^{\beta_0^\top \mathbf{Z}_i} ds \\
&\quad + \sum_{i=1}^n \int_0^t \frac{1}{\sqrt{n}} \left\{ 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}} - \frac{S_{j,\mathbf{z}}^{(0)}(s, \beta_0)}{S_j^{(0)}(s, \beta_0)} \right\} 1_{\{S_i=j\}} dM_i(s).
\end{aligned}$$

The first term in the right-hand side of this equality is 0 (straightforward calculations are omitted), which yields:

$$Q_{n,\mathbf{z}}^j(t, \beta_0) := \sum_{i=1}^n \int_0^t \frac{1}{\sqrt{n}} \left\{ 1_{\{\mathbf{Z}_i \leq \mathbf{z}\}} - \frac{S_{j,\mathbf{z}}^{(0)}(s, \beta_0)}{S_j^{(0)}(s, \beta_0)} \right\} 1_{\{S_i=j\}} dM_i(s). \quad (6)$$

The process $Q_{n,\mathbf{z}}^j := (Q_{n,\mathbf{z}}^j(\cdot, \beta_0))$ is then a martingale with respect to the filtration $\bigvee_{i \leq n} \mathcal{F}_{t,i}$. Let $k, \ell \in \{1, \dots, Q\}$ and $t \in \mathcal{T}$. Using expression (6), we have:

$$\begin{aligned}
\langle Q_{n,\mathbf{z}_k}^j, Q_{n,\mathbf{z}_\ell}^j \rangle(t) &= \sum_{i=1}^n \int_0^t \frac{1}{n} \left\{ 1_{\{\mathbf{Z}_i \leq \mathbf{z}_k\}} - \frac{S_{j,\mathbf{z}_k}^{(0)}(s, \beta_0)}{S_j^{(0)}(s, \beta_0)} \right\} \\
&\quad \times \left\{ 1_{\{\mathbf{Z}_i \leq \mathbf{z}_\ell\}} - \frac{S_{j,\mathbf{z}_\ell}^{(0)}(s, \beta_0)}{S_j^{(0)}(s, \beta_0)} \right\} 1_{\{S_i=j\}} \alpha_{0,j}(s) Y_i(s) e^{\beta_0^\top \mathbf{Z}_i} ds.
\end{aligned}$$

As n tends to infinity, $\langle Q_{n,\mathbf{z}_k}^j, Q_{n,\mathbf{z}_\ell}^j \rangle(t)$ converges in probability to

$$\begin{aligned}
C_{\mathbf{z}_k, \mathbf{z}_\ell}^j(t) &:= \mathbb{E} \left[\int_0^t 1_{\{\mathbf{Z} \leq \mathbf{z}_k\}} 1_{\{\mathbf{Z} \leq \mathbf{z}_\ell\}} 1_{\{S=j\}} \alpha_{0,j}(s) Y(s) e^{\beta_0^\top \mathbf{Z}} ds \right] \\
&\quad - \mathbb{E} \left[\int_0^t 1_{\{\mathbf{Z} \leq \mathbf{z}_k\}} \frac{s_{j,\mathbf{z}_\ell}^{(0)}(s, \beta_0)}{s_j^{(0)}(s, \beta_0)} 1_{\{S=j\}} \alpha_{0,j}(s) Y(s) e^{\beta_0^\top \mathbf{Z}} ds \right] \\
&\quad - \mathbb{E} \left[\int_0^t 1_{\{\mathbf{Z} \leq \mathbf{z}_\ell\}} \frac{s_{j,\mathbf{z}_k}^{(0)}(s, \beta_0)}{s_j^{(0)}(s, \beta_0)} 1_{\{S=j\}} \alpha_{0,j}(s) Y(s) e^{\beta_0^\top \mathbf{Z}} ds \right] \\
&\quad + \mathbb{E} \left[\int_0^t \frac{s_{j,\mathbf{z}_k}^{(0)}(s, \beta_0) s_{j,\mathbf{z}_\ell}^{(0)}(s, \beta_0)}{(s_j^{(0)}(s, \beta_0))^2} 1_{\{S=j\}} \alpha_{0,j}(s) Y(s) e^{\beta_0^\top \mathbf{Z}} ds \right] \\
&= \mathbb{E} \left[\int_0^t 1_{\{\mathbf{Z} \leq \mathbf{z}_k\}} 1_{\{\mathbf{Z} \leq \mathbf{z}_\ell\}} 1_{\{S=j\}} \alpha_{0,j}(s) Y(s) e^{\beta_0^\top \mathbf{Z}} ds \right] \\
&\quad - \int_0^t \frac{s_{j,\mathbf{z}_k}^{(0)}(s, \beta_0) s_{j,\mathbf{z}_\ell}^{(0)}(s, \beta_0)}{s_j^{(0)}(s, \beta_0)} \alpha_{0,j}(s) ds.
\end{aligned}$$

Moreover, when $j \neq j'$, we have:

$$\begin{aligned}
\langle Q_{n,\mathbf{z}_k}^j, Q_{n,\mathbf{z}_\ell}^{j'} \rangle(t) &= \sum_{i=1}^n \int_0^t \frac{1}{n} \left\{ 1_{\{\mathbf{Z}_i \leq \mathbf{z}_k\}} - \frac{S_{j,\mathbf{z}_k}^{(0)}(s, \beta_0)}{S_j^{(0)}(s, \beta_0)} \right\} \left\{ 1_{\{\mathbf{Z}_i \leq \mathbf{z}_\ell\}} - \frac{S_{j',\mathbf{z}_\ell}^{(0)}(s, \beta_0)}{S_{j'}^{(0)}(s, \beta_0)} \right\} \\
&\quad \times 1_{\{S_i=j\}} 1_{\{S_i=j'\}} Y_i(s) \sum_{m=1}^J \alpha_{0,m}(s) e^{\beta_0^\top \mathbf{Z}_i} 1_{\{S_i=m\}} ds \\
&= 0
\end{aligned}$$

since $1_{\{S_i=j\}} 1_{\{S_i=j'\}} = 0$. Next, we verify Lindeberg condition (*e.g.*, condition (3.18) of Theorem 5.3.5 in [7]). Let $\epsilon > 0$, $E_{j,\mathbf{z}_k}^{(0)}(s, \beta_0) := S_{j,\mathbf{z}_k}^{(0)}(s, \beta_0)/S_j^{(0)}(s, \beta_0)$ and define the jump process $Q_{n,\mathbf{z}_k,\epsilon}^j$ by $Q_{n,\mathbf{z}_k,\epsilon}^j(t, \beta_0) =$

$$\sum_{i=1}^n \int_0^t \frac{1}{\sqrt{n}} \left\{ 1_{\{\mathbf{Z}_i \leq \mathbf{z}_k\}} - E_{j,\mathbf{z}_k}^{(0)}(s, \beta_0) \right\} 1_{\left\{ \left| \frac{1}{\sqrt{n}} (1_{\{\mathbf{Z}_i \leq \mathbf{z}_k\}} - E_{j,\mathbf{z}_k}^{(0)}(s, \beta_0)) \right| > \epsilon \right\}} 1_{\{S_i=j\}} dM_i(s).$$

We have:

$$\begin{aligned}
\langle Q_{n,\mathbf{z}_k,\epsilon}^j, Q_{n,\mathbf{z}_k,\epsilon}^j \rangle(t) &= \sum_{i=1}^n \int_0^t \left\{ \frac{1}{\sqrt{n}} (1_{\{\mathbf{Z}_i \leq \mathbf{z}_k\}} - E_{j,\mathbf{z}_k}^{(0)}(s, \beta_0)) \right\}^2 1_{\{S_i=j\}} \\
&\quad \times \alpha_{0,j}(s) Y_i(s) e^{\beta_0^\top \mathbf{Z}_i} 1_{\left\{ \left| \frac{1}{\sqrt{n}} (1_{\{\mathbf{Z}_i \leq \mathbf{z}_k\}} - E_{j,\mathbf{z}_k}^{(0)}(s, \beta_0)) \right| > \epsilon \right\}} ds.
\end{aligned}$$

By using the inequality $|u - v|^2 1_{\{|u-v|>\epsilon\}} \leq 4u^2 1_{\{|u|>\epsilon/2\}} + 4v^2 1_{\{|v|>\epsilon/2\}}$, we obtain that $\langle Q_{n,\mathbf{z}_k,\epsilon}^j, Q_{n,\mathbf{z}_k,\epsilon}^j \rangle(t)$ is bounded above by

$$\begin{aligned} & \frac{4}{n} \sum_{i=1}^n \int_0^t 1_{\{\mathbf{z}_i \leq \mathbf{z}_k\}} 1_{\{1_{\{\mathbf{z}_i \leq \mathbf{z}_k\}} > \epsilon\sqrt{n}/2\}} Y_i(s) e^{\beta_0^\top \mathbf{z}_i} 1_{\{S_i=j\}} \alpha_{0,j}(s) ds \\ & + \frac{4}{n} \int_0^t \left(E_{j,\mathbf{z}_k}^{(0)}(s, \beta_0) \right)^2 1_{\{|E_{j,\mathbf{z}_k}^{(0)}(s, \beta_0)| > \epsilon\sqrt{n}/2\}} S_j^{(0)}(s, \beta_0) \alpha_{0,j}(s) ds. \end{aligned}$$

When n is sufficiently large, $1_{\{1_{\{\mathbf{z}_i \leq \mathbf{z}_k\}} > \epsilon\sqrt{n}/2\}} = 0$ for every $i = 1, \dots, n$ and the first term is zero. The second term converges to zero in probability, as the event $|E_{j,\mathbf{z}_k}^{(0)}(s, \beta_0)| > \epsilon\sqrt{n}/2$ cannot be true for large n under conditions C1-C5. By Theorem 5.3.5 in [7], the process $(Q_n^1(\cdot, \beta_0)^\top, \dots, Q_n^J(\cdot, \beta_0)^\top)$ converges weakly in $(D(\mathcal{T}))^{J \times Q}$ to a zero-mean Gaussian process $(\mathbb{W}^{1\top}, \dots, \mathbb{W}^{J\top})$ (with $\mathbb{W}^j := (W_1^j, \dots, W_Q^j)^\top$) with covariance function $\text{cov}(W_k^j(t_1), W_\ell^j(t_2)) = C_{\mathbf{z}_k, \mathbf{z}_\ell}^j(\min(t_1, t_2))$ and $\text{cov}(W_k^j(t_1), W_\ell^{j'}(t_2)) = 0$ if $j \neq j'$, for $k, \ell = 1, \dots, Q$.

□

Proof of Lemma 2 By a first-order Taylor expansion of $\beta \mapsto Q_{n,\mathbf{z}}^j(t, \beta)$ around β_0 , we have:

$$Q_{n,\mathbf{z}}^j(t, \hat{\beta}_n) - Q_{n,\mathbf{z}}^j(t, \beta_0) = (\hat{\beta}_n - \beta_0)^\top \frac{\partial}{\partial \beta} Q_{n,\mathbf{z}}^j(t, \tilde{\beta}_n),$$

where $\tilde{\beta}_n$ is on the line segment between $\hat{\beta}_n$ and β_0 . The derivative $\partial Q_{n,\mathbf{z}}^j(t, \beta) / \partial \beta$ is given by:

$$\frac{\partial}{\partial \beta} Q_{n,\mathbf{z}}^j(t, \beta) = -\frac{1}{\sqrt{n}} \int_0^t H_{j,\mathbf{z}}(s, \beta) d\bar{M}_j(s),$$

where $\bar{M}_j(s) = \sum_{i=1}^n 1_{\{S_i=j\}} M_i(s)$ and

$$H_{j,\mathbf{z}}(s, \beta) := \frac{S_{j,\mathbf{z}}^{(1)}(s, \beta)}{S_j^{(0)}(s, \beta)} - \frac{S_{j,\mathbf{z}}^{(0)}(s, \beta) S_j^{(1)}(s, \beta)}{(S_j^{(0)}(s, \beta))^2}.$$

Then, by Cauchy-Schwarz inequality,

$$\left\| Q_{n,\mathbf{z}}^j(\cdot, \hat{\beta}_n) - Q_{n,\mathbf{z}}^j(\cdot, \beta_0) \right\|_{\mathcal{T}} \leq \|\hat{\beta}_n - \beta_0\| \left\| \frac{\partial}{\partial \beta} Q_{n,\mathbf{z}}^j(\cdot, \tilde{\beta}_n) \right\|_{\mathcal{T}}.$$

Under H_0 , $\hat{\beta}_n$ converges in probability to β_0 as n tends to infinity and thus, $\|\hat{\beta}_n - \beta_0\|$ converges to zero. We prove that $\left\| \frac{\partial}{\partial \beta} Q_{n,\mathbf{z}}^j(\cdot, \tilde{\beta}_n) \right\|_{\mathcal{T}}$ is bounded in probability. Let

$$h_{j,\mathbf{z}}(s, \beta) := \frac{s_{j,\mathbf{z}}^{(1)}(s, \beta)}{s_j^{(0)}(s, \beta)} - \frac{s_{j,\mathbf{z}}^{(0)}(s, \beta) s_j^{(1)}(s, \beta)}{(s_j^{(0)}(s, \beta))^2}.$$

We prove that $\|H_{j,\mathbf{z}}(\cdot, \tilde{\beta}_n) - h_{j,\mathbf{z}}(\cdot, \beta_0)\|_{\mathcal{T}}$ converges in probability to zero as n tends to infinity. We have:

$$\begin{aligned} \|H_{j,\mathbf{z}}(\cdot, \tilde{\beta}_n) - h_{j,\mathbf{z}}(\cdot, \beta_0)\|_{\mathcal{T}} &\leq \left\| \frac{n^{-1}S_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)}{n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n)} - \frac{s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)}{s_j^{(0)}(\cdot, \beta_0)} \right\|_{\mathcal{T}} \\ &\quad + \left\| \frac{s_{j,\mathbf{z}}^{(0)}(\cdot, \beta_0)s_j^{(1)}(\cdot, \beta_0)}{(s_j^{(0)}(\cdot, \beta_0))^2} - \frac{S_{j,\mathbf{z}}^{(0)}(\cdot, \tilde{\beta}_n)S_j^{(1)}(\cdot, \tilde{\beta}_n)}{(S_j^{(0)}(\cdot, \tilde{\beta}_n))^2} \right\|_{\mathcal{T}} \\ &:= U_{n,1} + U_{n,2}. \end{aligned}$$

We prove that $U_{n,1}$ converges in probability to zero as n tends to infinity (arguments are similar for $U_{n,2}$ and are thus omitted). Using the elementary equality $\frac{u_n}{v_n} - \frac{u}{v} = (\frac{u_n}{v} - \frac{uv_n}{v^2})\frac{v}{v_n}$, we obtain:

$$\begin{aligned} U_{n,1} &= \left\| \left(\frac{n^{-1}S_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)}{s_j^{(0)}(\cdot, \beta_0)} - \frac{s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n)}{(s_j^{(0)}(\cdot, \beta_0))^2} \right) \frac{s_j^{(0)}(\cdot, \beta_0)}{n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n)} \right\|_{\mathcal{T}} \\ &\leq \left\| \frac{n^{-1}S_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)}{s_j^{(0)}(\cdot, \beta_0)} - \frac{s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n)}{(s_j^{(0)}(\cdot, \beta_0))^2} \right\|_{\mathcal{T}} \left\| \frac{s_j^{(0)}(\cdot, \beta_0)}{n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n)} \right\|_{\mathcal{T}} \\ &\leq \frac{1}{c^2} \left\| n^{-1}S_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)s_j^{(0)}(\cdot, \beta_0) - s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n) \right\|_{\mathcal{T}} \left\| \frac{s_j^{(0)}(\cdot, \beta_0)}{n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n)} \right\|_{\mathcal{T}} \end{aligned}$$

where the second to third line follows by condition C5. Now,

$$\begin{aligned} &\left\| n^{-1}S_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)s_j^{(0)}(\cdot, \beta_0) - s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n) \right\|_{\mathcal{T}} \\ &\leq \left\| n^{-1}S_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)s_j^{(0)}(\cdot, \beta_0) - s_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)s_j^{(0)}(\cdot, \beta_0) \right\|_{\mathcal{T}} \\ &\quad + \left\| s_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)s_j^{(0)}(\cdot, \beta_0) - s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)s_j^{(0)}(\cdot, \beta_0) \right\|_{\mathcal{T}} \\ &\quad + \left\| s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)s_j^{(0)}(\cdot, \beta_0) - s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)s_j^{(0)}(\cdot, \tilde{\beta}_n) \right\|_{\mathcal{T}} \\ &\quad + \left\| s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)s_j^{(0)}(\cdot, \tilde{\beta}_n) - s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n) \right\|_{\mathcal{T}} \end{aligned}$$

and thus

$$\begin{aligned} &\left\| n^{-1}S_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)s_j^{(0)}(\cdot, \beta_0) - s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n) \right\|_{\mathcal{T}} \\ &\leq \left\| s_j^{(0)}(\cdot, \beta_0) \right\|_{\mathcal{T}} \left(\left\| n^{-1}S_{j,\mathbf{z}}^{(1)} - s_{j,\mathbf{z}}^{(1)} \right\|_{\mathcal{T} \times \mathcal{B}} + \left\| s_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n) - s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0) \right\|_{\mathcal{T}} \right) \\ &\quad + \left\| s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0) \right\|_{\mathcal{T}} \left(\left\| s_j^{(0)}(\cdot, \beta_0) - s_j^{(0)}(\cdot, \tilde{\beta}_n) \right\|_{\mathcal{T}} + \left\| s_j^{(0)} - n^{-1}S_j^{(0)} \right\|_{\mathcal{T} \times \mathcal{B}} \right). \end{aligned}$$

From Glivenko-Cantelli theorem, $\|n^{-1}S_{j,\mathbf{z}}^{(1)} - s_{j,\mathbf{z}}^{(1)}\|_{\mathcal{T} \times \mathcal{B}}$ and $\|n^{-1}S_j^{(0)} - s_j^{(0)}\|_{\mathcal{T} \times \mathcal{B}}$ converge to zero. By condition C4, $\|s_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n) - s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)\|_{\mathcal{T}}$ and $\|s_j^{(0)}(\cdot, \beta_0) - s_j^{(0)}(\cdot, \tilde{\beta}_n)\|_{\mathcal{T}}$ converge to zero. Moreover, $\|s_j^{(0)}(\cdot, \beta_0)\|_{\mathcal{T}}$ and $\|s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)\|_{\mathcal{T}}$ are bounded by conditions C1 and C3. Thus, the quantity $\|n^{-1}S_{j,\mathbf{z}}^{(1)}(\cdot, \tilde{\beta}_n)s_j^{(0)}(\cdot, \beta_0) - s_{j,\mathbf{z}}^{(1)}(\cdot, \beta_0)n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n)\|_{\mathcal{T}}$ tends to zero. Using similar arguments, one can show that $\|s_j^{(0)}(\cdot, \beta_0)/n^{-1}S_j^{(0)}(\cdot, \tilde{\beta}_n)\|_{\mathcal{T}}$ converges to 1 and finally, $U_{n,1}$ converges to zero. Convergence to zero of $U_{n,1}$ and $U_{n,2}$ implies that $\|H_{j,\mathbf{z}}(\cdot, \tilde{\beta}_n) - h_{j,\mathbf{z}}(\cdot, \beta_0)\|_{\mathcal{T}}$ converges in probability to zero. Now, we have:

$$\|H_{j,\mathbf{z}}(\cdot, \tilde{\beta}_n)\|_{\mathcal{T}} \leq \|H_{j,\mathbf{z}}(\cdot, \tilde{\beta}_n) - h_{j,\mathbf{z}}(\cdot, \beta_0)\|_{\mathcal{T}} + \|h_{j,\mathbf{z}}(\cdot, \beta_0)\|_{\mathcal{T}}.$$

Under conditions C1, C3, C5, $\|h_{j,\mathbf{z}}(\cdot, \beta_0)\|_{\mathcal{T}}$ is bounded and thus $\|H_{j,\mathbf{z}}(\cdot, \tilde{\beta}_n)\|_{\mathcal{T}}$ and then $\|\frac{\partial}{\partial \beta} Q_{n,\mathbf{z}}^j(\cdot, \tilde{\beta}_n)\|_{\mathcal{T}}$ are bounded in probability. Finally, $Q_{n,\mathbf{z}}^j(\cdot, \tilde{\beta}_n) = Q_{n,\mathbf{z}}^j(\cdot, \beta_0) + o_{\mathbb{P}}(1)$ where the $o_{\mathbb{P}}(1)$ is uniform on \mathcal{T} . This concludes the proof.

□

Appendix B. Parameters values in the simulation study for $J = 5$.

For $J = 5$, we consider the following alternatives (respectively denoted by $H_{1,a}, H_{1,b}, H_{1,c}$ and $H_{1,d}$ in Table 2):

- *non-proportional hazards model*: the hazard function in stratum j is chosen as $\alpha_j(t) = \alpha_j e^{1.2Z_1 + 1.5Z_2 \times t}$, $j = 1, \dots, 5$. We take $\alpha_1 = 0.01, \alpha_2 = 0.1, \alpha_3 = 0.25, \alpha_4 = 0.3, \alpha_5 = 0.2$.
- *covariate threshold effect*: the hazard function in stratum j is chosen as $\alpha_j(t) = \lambda_j \alpha_j t^{\alpha_j - 1} e^{1.7Z_1 1_{\{Z_1 > \xi_j\}} + 0.5Z_2}$, $j = 1, \dots, 5$. We take $(\alpha_1, \lambda_1, \xi_1) = (1.5, 1, 0.6)$, $(\alpha_2, \lambda_2, \xi_2) = (0.5, 0.75, 1)$, $(\alpha_3, \lambda_3, \xi_3) = (1, 1.25, 0.8)$, $(\alpha_4, \lambda_4, \xi_4) = (0.75, 0.8, 1.2)$ and $(\alpha_5, \lambda_5, \xi_5) = (1, 0.8, 0.75)$.
- *distinct regression parameters across strata*: the hazard function in stratum j is chosen as $\alpha_j(t) = \lambda_j \alpha_j t^{\alpha_j - 1} e^{\beta_{1,j}Z_1 + \beta_{2,j}Z_2}$, $j = 1, \dots, 5$. We take $(\alpha_1, \lambda_1, \beta_{1,1}, \beta_{2,1}) = (2.1, 1, 0.2, 0.7)$, $(\alpha_2, \lambda_2, \beta_{1,2}, \beta_{2,2}) = (1.2, 0.75, 1, 1)$, $(\alpha_3, \lambda_3, \beta_{1,3}, \beta_{2,3}) = (1.8, 1.5, 0.2, 0.2)$, $(\alpha_4, \lambda_4, \beta_{1,4}, \beta_{2,4}) = (1, 1, 1.3, 0.5)$ and $(\alpha_5, \lambda_5, \beta_{1,5}, \beta_{2,5}) = (1.5, 0.5, 0.15, 0.55)$.
- *omitted covariates*: the hazard function in stratum j is taken as $\alpha_j(t) = \lambda_j \alpha_j t^{\alpha_j - 1} e^{Z_1 - 0.7Z_2 + 0.75Z_3}$ ($j = 1, \dots, 5$) where the additional covariate Z_3 is distributed as a $\mathcal{N}(1, 0.25)$. We take $(\alpha_1, \lambda_1) = (2.1, 1)$, $(\alpha_2, \lambda_2) = (1.2, 0.75)$, $(\alpha_3, \lambda_3) = (1.8, 1.5)$, $(\alpha_4, \lambda_4) = (1.2, 1.25)$ and $(\alpha_5, \lambda_5) = (0.5, 0.8)$.

Acknowledgements

The authors acknowledge financial support from Monastir University (Tunisia) and INSA of Rennes (France). Part of this work was made while R. Ben Elouefi was visiting INSA of Rennes and while J.-F. Dupuy was visiting Monastir University.

References

- [1] P. K. Andersen, Ø. Borgan, R. D. Gill, and N. Keiding. *Statistical models based on counting processes*. Springer Series in Statistics. Springer-Verlag, New York, 1993.
- [2] V. Bagdonavičius and M. S. Nikulin. *Accelerated life models: modeling and statistical analysis*. Boca Raton, FL: CRC Press, 2002.
- [3] The Fibrinogen Studies Collaboration. Measures to assess the prognostic ability of the stratified Cox proportional hazards model. *Stat. Med.*, 28(3):389–411, 2009.
- [4] D. R. Cox. Regression models and life-tables. *J. Roy. Statist. Soc. Ser. B*, 34:187–220, 1972. With discussion.
- [5] D. R. Cox. Partial likelihood. *Biometrika*, 62(2):269–276, 1975.
- [6] J.-F. Dupuy and E. Leconte. A study of regression calibration in a partially observed stratified Cox model. *J. Statist. Plann. Inference*, 139(2):317–328, 2009.
- [7] T. R. Fleming and D. P. Harrington. *Counting Processes and Survival Analysis*. Wiley Series in Probability and Statistics. Wiley, 1991.
- [8] A. Gandy and U. Jensen. On goodness-of-fit tests for Aalen’s additive risk model. *Scand. J. Statist.*, 32(3):425–445, 2005.
- [9] A. Gandy and U. Jensen. Model checks for Cox-type regression models based on optimally weighted martingale residuals. *Lifetime Data Anal.*, 15(4):534–557, 2009.
- [10] M. Gorfine, L. Hsu, and R. L. Prentice. Nonparametric correction for covariate measurement error in a stratified Cox model. *Biostatistics*, 5(1):75–87, 2004.
- [11] M. Haenlein. Social interactions in customer churn decisions: The impact of relationship directionality. *Int. J. Research Marketing*, 30(3):236–248, 2013.

- [12] T. E. Hanson, A. Jara, and L. Zhao. A Bayesian semiparametric temporally-stratified proportional hazards model with spatial frailties. *Bayesian Anal.*, 7(1):147–188, 2012.
- [13] H. Heinzl, J. Stare, and M. Mittlböck. A measure of dependence for the stratified Cox proportional hazards regression model. *Biom. J.*, 44(6):671–682, 2002.
- [14] N. M. Kiefer and C. E. Larson. Counting processes for retail default modeling. Technical Report CREATES Research Paper 2015-17, Department of Economics and Business Economics, Aarhus University, Denmark, 2015.
- [15] J. Kim. Confidence intervals for the difference of median survival times using the stratified Cox proportional hazards model. *Biom. J.*, 43, 2001.
- [16] D. Y. Lin, L. J. Wei, and Z. Ying. Checking the Cox model with cumulative sums of martingale-based residuals. *Biometrika*, 80(3):557–572, 1993.
- [17] T. Martinussen and T. H. Scheike. *Dynamic regression models for survival data*. Statistics for Biology and Health. Springer, New York, 2006.
- [18] L. Marzec and P. Marzec. Generalized martingale-residual processes for goodness-of-fit inference in Cox’s type regression models. *Ann. Statist.*, 25(2):683–714, 1997.
- [19] L. Natarajan and J. O’Quigley. Predictive capability of stratified proportional hazards models. *J. Appl. Stat.*, 29(8):1153–1163, 2002.
- [20] A. K. Örtqvist, C. Lundholm, H. Kieler, J. F. Ludvigsson, T. Fall, W. Ye, and C. Almqvist. Antibiotics in fetal and early life and subsequent childhood asthma: nationwide population based study with sibling analysis. *BMJ*, 349:doi: 10.1136/bmj.g6979, 2014.
- [21] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014.
- [22] F. Scavonetto, T. Y. Yeoh, E. C. Umbreit, T. N. Weingarten, M. T. Gettman, I. Frank, S. A. Boorjian, R. J. Karnes, D. R. Schroeder, L. J. Rangel, A. C. Hanson, R. E. Hofer, D. I. Sessler, and J. Sprung. Association between neuraxial analgesia, cancer progression, and mortality after radical prostatectomy: a large, retrospective matched cohort study. *Br. J. Anaesth.*, 113:doi: 10.1093/bja/aet467, 2014.
- [23] W. Stute and L.-X. Zhu. Model checks for generalized linear models. *Scand. J. Statist.*, 29(3):535–545, 2002.
- [24] J. Q. Su and L. J. Wei. A lack-of-fit test for the mean function in a generalized linear model. *J. Amer. Statist. Assoc.*, 86(414):420–426, 1991.

- [25] P. J. M. Verweij, H. C. van Houwelingen, and T. Stijnen. A goodness-of-fit test for Cox's proportional hazards model based on martingale residuals. *Biometrics*, 54(4):1517–1526, 1998.