

Research Article

Analysis of Heart Transplant Survival Data Using Generalized Additive Models

Masaaki Tsujitani¹ and Yusuke Tanaka²

¹ Department of Engineering Informatics, Osaka Electro-Communication University, Osaka 572-8530, Japan

² Clinical Information Division Data Science Center, EPS Corporation, Osaka 532-0003, Japan

Correspondence should be addressed to Masaaki Tsujitani; ekaaf900@ricv.zaq.ne.jp

Received 13 March 2013; Accepted 7 May 2013

Academic Editor: Frank Emmert-Streib

Copyright © 2013 M. Tsujitani and Y. Tanaka. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The Stanford Heart Transplant data were collected to model survival in patients using penalized smoothing splines for covariates whose values change over the course of the study. The basic idea of the present study is to use a logistic regression model and a generalized additive model with B -splines to estimate the survival function. We model survival time as a function of patient covariates and transplant status and compare the results obtained using smoothing spline, partial logistic, Cox's proportional hazards, and piecewise exponential models.

1. Introduction

Cox's proportional hazards model has been proposed based on the relationship between survival and the patient characteristics observed when the patient entered the study [1]. When the values of covariates change over the course of the study, however, a number of theoretical problems with respect to the baseline survival function and the baseline cumulative hazard function need to be solved [2]. Several prognostic models [3–6] have become as widely used as Cox's proportional hazards model for the analysis of survival data having time-dependent covariates. The present study examines the nonlinear effects of the evolution of the covariates over time using penalized smoothing splines.

Cox's proportional hazards model postulates that the hazard at time t is the product of two components:

$$h(t) = h_0(t) \exp \left[\sum_{i=1}^I b_i x_i \right], \quad (1)$$

where $\mathbf{b} = (b_1, \dots, b_I)$ is a vector of coefficients. The proportional hazards assumption is that the baseline hazard $h_0(t)$ is a function of t but does not involve the values of covariates. Several prognostic models for heart transplant survival data have been developed using Cox's regression analysis, and

the values of all covariates are determined at the time when the patient entered the study [7–9]. However, situations may exist in which the values of covariates change over the course of the study. A time-dependent model uses the follow-up data to estimate the effect of the evolution of the covariates over time. The relative hazard $h(t)/h_0(t)$ then depends on time t , and thus the proportional hazards assumption is no longer satisfied [6, 10].

The time-dependent covariates $\mathbf{X}_l^{(d)} = (x_{l1}^{(d)}, \dots, x_{lI}^{(d)})$ are provided for patient no. d , where $x_{l1}^{(d)}$ is the midpoint of the l th time interval. Given the continuous survivor time, piecewise models arise from the partition of the time axis into disjointed intervals. Biganzoli et al. [11, 12] show that by treating the time interval as an input variable in a feed-forward neural network, it is possible to estimate smoothed discrete hazards as conditional probabilities of failure. To apply a generalized additive model (GAM), discretization of one-month or one-week intervals must be applied for the continuous survivor time with time-fixed covariates. However we cannot determine which discretization, one-month or one-week, should be applied; that is, the discretization is not initially unique. In the case of time-dependent covariates $\mathbf{X}_l^{(d)}$, $x_{l1}^{(d)}$ is initially determined as the midpoint of the l th time interval for patient no. d . It is fairly straightforward

to extend the model to survivor data with time-dependent covariates. Furthermore, by regarding a GAM as an extension of a partial logistic model (PLM), the unknown parameters can be estimated by maximizing the partial log-likelihood [13, 14].

We use the Stanford Heart Transplant data, which has been collected to model survival in patients. Although Cox's proportional hazards model is not applicable in the case of time-dependent covariates, the survivor function can be estimated by taking $h_0(t)$ in (1) to be the piecewise exponential hazard. Crowley and Hu [7], Aitkin et al. [8], and Lawless [9] used piecewise exponential models and plotted the survival function. Lagakos [15] also examined a graphical technique for assessing covariates in Cox's proportional hazards model based on a permutation of the observed rank statistic. Most previous studies compared the hazard functions to assess the effect of transplantation on survival by fitting pretransplant and posttransplant data separately.

The difficulty is that there is no easily used measure of the difference between the transplanted and nontransplanted groups. Inferences must be based on a comparison of the estimated function. As Aitkin et al. [8] pointed out, there are always dangers in making inferences about the effect of treatment without adequate control groups. We thus provide an analysis that includes pretransplant and posttransplant data simultaneously as time-dependent covariates. It should be emphasized that patients who are not transplanted constitute a control group relative to patients who have undergone heart transplantation by the same covariates.

We use the 1977 version of the data, as given in Crowley and Hu [7], which is for 103 patients. As four of the transplanted patients have incomplete data on the mismatch score, our analysis is based on 99 patients to assess for what values of these covariates, if any, transplantation is likely to prolong survival. More than 30 percent of cases are censored. In these data, survival times are the number of days until death following a heart transplant, as in Lagakos [15]. A distinctive feature of the present problem is that some of the covariates are time-dependent (and possibly random). For example, Table 1 shows the values of covariates for transplant status (i.e., waiting time), age at transplant (in years), mismatch score (as time-dependent covariates), and previous open-heart surgery for patient no. 18. The previous surgery status does not change with time. In order to extend this setting, the covariate for transplant status is taken as an indicator (coded as 0 before the point of transplant and 1 after transplant). All the other time-dependent covariates are treated as being zero before transplant but changing from zero to the actual value of the particular covariate at the time of transplant. Patient no. 18 generated six observations. The proposed methods allow for simultaneous investigation of several covariates and provide estimates of the survival function as well as the significance.

2. Generalized Additive Models

By extending the PLM for the grouped data based on partial likelihood as introduced by Cox [16] and Efron [17], a PLM

can be proposed for ungrouped data [13, 14] having time-dependent covariates for the discrete hazard rate $h_l^{(d)}$ of patient no. d at the time interval l :

$$\begin{aligned} \text{PLM: } \ln \left(\frac{h_l^{(d)}}{1 - h_l^{(d)}} \right) &= \beta_0 + \beta_1 x_{l1}^{(d)} \\ &+ \beta_2 x_{l2}^{(d)} + \beta_3 x_{l3}^{(d)} \\ &+ \cdots + \beta_I x_{lI}^{(d)}. \end{aligned} \quad (2)$$

In recent years, a variety of powerful techniques have been developed for exploring the functional form of effects. Here, GAM with smoothing splines proposed by Hastie et al. [18, 19] will be used by extending the generalized linear model (GLM) in McCullagh and Nelder [20], where the linear predictor in (2) is specified as a sum of smooth functions $s(x)$ with twice continuously differentiable functions of some or all of the covariates:

$$\begin{aligned} \text{GAM: } \ln \left(\frac{h_l^{(d)}}{1 - h_l^{(d)}} \right) &= \beta_0 + s_1(x_{l1}^{(d)}) \\ &+ s_2(x_{l2}^{(d)}) + s_3(x_{l3}^{(d)}) \\ &+ \cdots + s_I(x_{lI}^{(d)}). \end{aligned} \quad (3)$$

The smooth functions in (3) can be represented as

$$\begin{aligned} s_0(x) &= \sum_{j=1}^{q_0} \beta_j b_{0j}(x), \\ s_1(x) &= \sum_{j=1}^{q_1} \beta_{q_0+j} b_{1j}(x), \\ &\vdots \\ s_I(x) &= \sum_{j=1}^{q_I} \beta_{q_{I-1}+j} b_{Ij}(x), \end{aligned} \quad (4)$$

where q_1, q_2, \dots, q_I are the numbers of knots, and

$$\begin{aligned} \boldsymbol{\beta} &= (\beta_0, \beta_1, \dots, \beta_{q_0}, \beta_{q_0+1}, \beta_{q_0+2}, \dots, \beta_{q_0+q_1}, \\ &\dots, \beta_{q_0+q_1+\dots+q_{I-1}+1}, \beta_{q_0+q_1+\dots+q_{I-1}+2}, \\ &\dots, \beta_{q_0+q_1+\dots+q_I}). \end{aligned} \quad (5)$$

For time interval l of patient no. d , we have the following definitions:

$$\delta_l^{(d)} = \begin{cases} 1: \text{patient no. } d \text{ died at time interval} \\ \quad \text{for the } l\text{th clinic visit,} \\ 0: \text{otherwise,} \end{cases}$$

$$\delta_l^{\prime(d)} = \begin{cases} 1: \text{patient no. } d \text{ was censored at} \\ \quad \text{time interval } l, \\ 0: \text{o.w.,} \end{cases}$$

TABLE I: Covariate values for patient no. 18.

Time interval l	Midpoint x_1 (days)	Transplant status x_2	Age at transplant x_3 (years)	Mismatch score x_4	Previous surgery x_5
1	3.5	0	0	0	0
2	10.5	0	0	0	0
3	17.5	1	56	2.05	0
4	24.5	1	56	2.05	0
5	31.5	1	56	2.05	0
6	38.5	1	56	2.05	0

$$\begin{aligned}
\mathbf{v}_l^{(d)} &= (\delta_1^{(d)}, \delta_1'^{(d)}, \delta_2^{(d)}, \delta_2'^{(d)}, \\
&\quad \dots, \delta_{l-1}^{(d)}, \delta_{l-1}'^{(d)}) = (0, 0, \dots, 0), \\
\mathbf{v}_l'^{(d)} &= (\mathbf{v}_l^{(d)}, \delta_l^{(d)}) = (0, 0, \dots, 0, \delta_l^{(d)}),
\end{aligned} \tag{6}$$

where $\mathbf{v}_l^{(d)}$ is the history of defaults and is censored for the first $l-1$ time intervals of patient no. d and $\mathbf{v}_l'^{(d)} = (\mathbf{v}_l^{(d)}, \delta_l^{(d)})$ is the same history extended to include $\delta_l^{(d)}$. Using the above model and notation, Tsujitani and Sakon [13] derived the full log-likelihood for all patients

$$\ln L = \ln L(\boldsymbol{\beta}) + \sum_{d=1}^n \sum_{l=1}^{l_d} \ln p(\delta_l^{(d)} | \mathbf{v}_l'^{(d)}) \tag{7}$$

with the partial log-likelihood

$$\begin{aligned}
\ln L(\boldsymbol{\beta}) &= \sum_{d=1}^n \left\{ \sum_{l=1}^{l_d-1} \ln(1 - h_l^{(d)}) + \delta_{l_d}^{(d)} \ln h_{l_d}^{(d)} \right. \\
&\quad \left. + (1 - \delta_{l_d}^{(d)}) \ln(1 - h_{l_d}^{(d)}) \right\}.
\end{aligned} \tag{8}$$

Although $\ln L(\boldsymbol{\beta})$ is not a log-likelihood in the usual sense, it possesses the usual asymptotic properties under fairly broad conditions, as proven in Andersen and Gill [21]. To avoid overfitting, such models are estimated by penalized maximum likelihood

$$\begin{aligned}
\ln L(\boldsymbol{\beta}) &= \sum_{d=1}^n \left\{ \sum_{l=1}^{l_d-1} \ln(1 - h_l^{(d)}) \right. \\
&\quad \left. + \delta_{l_d}^{(d)} \ln h_{l_d}^{(d)} + (1 - \delta_{l_d}^{(d)}) \ln(1 - h_{l_d}^{(d)}) \right\} \\
&\quad + \frac{1}{2} \sum_{i=1}^I \lambda_i \int \{s_i''(t)\}^2 dt,
\end{aligned} \tag{9}$$

where λ_i are smoothing parameters that control the trade-off between the fit and the smoothness. The functions $s_i(x)$ in (9) are represented by the B -spline basis functions $b_i(x)$; see, for details, Tsujitani et al. [14].

Two model-fitting issues remain. The first concerns the selection of smoothing parameter λ_i in (9). The optimum smoothing parameter choice is outweighed by the easy identification of a covariate's functional form as well as the applicability of established inferential methods to short-term survival prediction. In order to select the smoothing parameters, the algorithm developed by Wood [22–24] can be applied by minimizing generalized cross validation (GCV) as an approximation to leave-one-out CV [23]. It should be noted that the leaving-one-out CV is to allow the deletion of only one observation. On the other hand, the ordinal ν -fold CV divides the data randomly into ν groups so that their sizes are as nearly equal as possible. This partition should be made to avoid possible biases, as described in Zhang [25]. In many problems, the ordinal ν -fold CV is, thus, unsatisfactory in several respects for time-dependent covariates. Applying this kind of data structure to the CV algorithm, we obtain insights into how the partition of data should be carried out. A natural extension of the ν -fold CV algorithm by setting $\nu = n$ is to allow the deletion of the patient with several observations; see, for details, Tsujitani et al. [14].

A second issue is the goodness-of-fit test of the model. After choosing the optimum smoothing parameters via the variant ν -fold CV algorithm, the deviance allows us to test the goodness-of-fit:

$$\text{Dev} = 2(\ln L_{\max} - \ln L_c), \tag{10}$$

where $\ln L_c$ denotes the maximized partial log-likelihood under some current GAM and the log-likelihood for the maximum (full) model $\ln L_{\max}$ is zero. The deviance (10) is, however, not even approximately a χ^2 distribution for the case in which ungrouped binary responses are available; see, for example, Collett [26], Landwehr et al. [27], and Tsujitani and Sakon [13]. The number of degrees of freedom required for the test for significance using the assumed χ^2 distribution for the deviance is a contentious issue. No adequate distribution theory yet exists for the deviance. The reason for this is somewhat technical; for details, see Section 3.8 in Collett [2]. Consequently, the deviance on fitting a model to binary response data cannot be used as a summary measure of the goodness-of-fit of the model. Thus, bootstrapping is applied to the deviance (10) in order to obtain the goodness-of-fit; see, for details, Efron and Tibshirani [28] and Tsujitani et al. [14].

TABLE 2: Optimum smoothing parameters.

Covariates	Variant ν -fold CV	GCV
Midpoint (x_1)	0.0001	2.75×10^{-12}
Age (x_3)	0.01	1.30×10^{-6}
Mismatch score (x_4)	0.01	2.53×10^{-6}

3. Example

As an initial model for the Stanford Heart Transplant data, we employ

$$\text{Model I: } s(x_1) + x_2 + s(x_3) + s(x_4) + x_5. \quad (11)$$

GCV is only an approximation of leaving-one-out CV. Alternatively the variant ν -fold CV is leaving-one-out CV based on each of $n = 99$ patients to allow the deletion of the patient with several observations. By using variant ν -fold CV and GCV for the initial model, the optimum smoothing parameters for GAM are determined as shown in Table 2. By using a backward elimination procedure, we obtain

$$\text{Model II: } s(x_1) + x_2 + s(x_3) + x_5. \quad (12)$$

The likelihood ratio (LR) statistic based on deviance can be computed to test the significance of spline effects (i.e., nonlinearity). For example, the spline effect of “Midpoint(x_1)” can also be tested by using

$$\text{Model III: } x_1 + x_2 + s(x_3) + x_5. \quad (13)$$

By comparing Model II with III, the reduction in the value of deviance is $\Delta = 4.59$ with $1.85d.f.$ This is significant at the 10% level. The spline effect for “Age(x_3)” is not significant. We thus obtain the final optimum GAM

$$\text{Model IV: } s(x_1) + x_2 + x_3 + x_5 \quad (14)$$

with a variant ν -fold score of 654.754.

Figure 1 shows a histogram of the bootstrapped $\text{Dev}(b)$ for the optimum model. The bootstrap estimate of the 95th percentile Dev^* is $\text{Dev}^* = 685.65$. The comparison to $\text{Dev} = 639.77$ of (10) suggests that the model fits the data.

Figure 2 shows the estimated contribution $\hat{s}(x_1)$ of “Midpoint(x_1)” to $\ln\{\hat{h}_1^{(d)} / (1 - \hat{h}_1^{(d)})\}$, together with the ± 2 standard deviation (SD) curves for the final optimum Model IV. The spline effects of x_1 are visualized in Figure 2. Figure 2 nicely shows that the spline function $s(x_1)$ of dying decreases initially as the midpoint x_1 increases. Subsequently, however, $s(x_1)$ is stably maintained after 500 midpoint. For the purpose of comparison, Figure 3 shows the estimated contribution $\hat{s}(x_1)$ for GCV. From Figure 3, it is clear that the estimated $\hat{s}(x_1)$ of x_1 is flat until 1500 and then tumbles because of too small smoothing parameter (i.e., overfitting), as shown in Table 2. So variant ν -fold CV is superior to GCV. The analyses in this example are carried out using library `{mgcv}` in R.

The survival function for our discretized situation is

$$\Pr(l) = \prod_{1 \leq i \leq l} (1 - h_i^{(d)}). \quad (15)$$

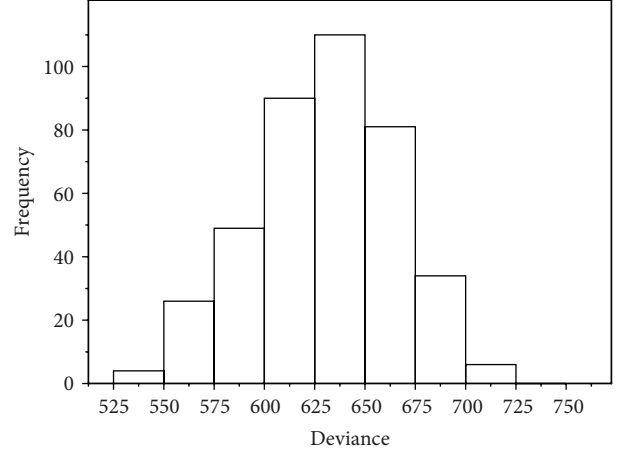
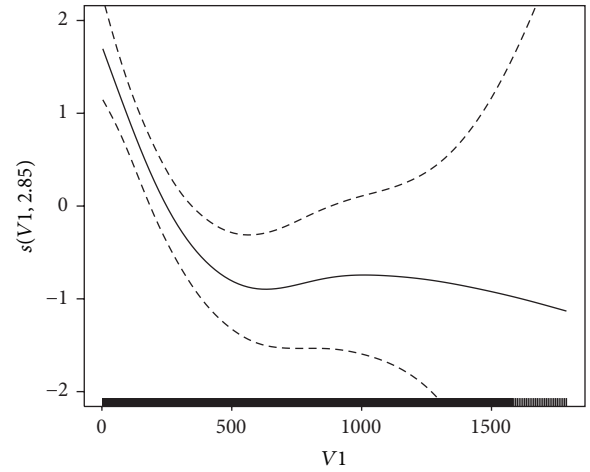
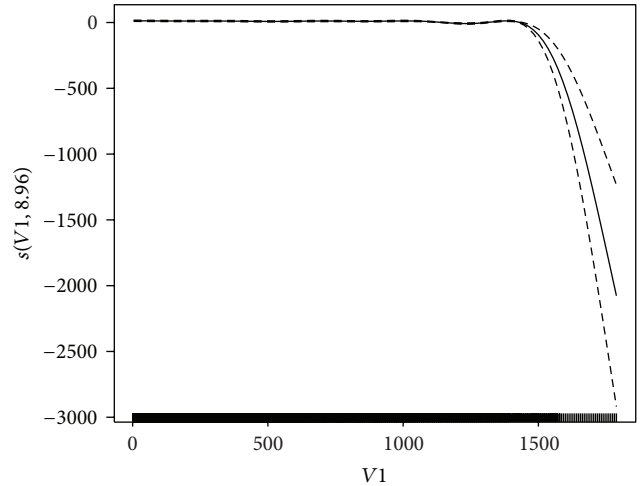
FIGURE 1: Histogram of the bootstrapped $\text{Dev}(b)$ for $B = 400$.FIGURE 2: Estimated contribution $\hat{s}(x_1)$ of x_1 (solid curve) and $\hat{s}(x_1) \pm 2\text{SD}[\hat{s}(x_1)]$ (dashed curves).FIGURE 3: Estimated contribution $\hat{s}(x_1)$ of x_1 (solid curve) and $\hat{s}(x_1) \pm 2\text{SD}[\hat{s}(x_1)]$ (dashed curves) for GCV.

TABLE 3: P values for the significance test of covariates.

Covariates	GAM	Partial logistic	Proportional hazard	Piecewise exponential
Transplant status (x_2)	0.0107	<0.0001	0.0076	0.0081
Age (x_3)	0.011	0.0135	0.0190	0.0199
Previous surgery (x_5)	0.0575	0.0672	0.0830	0.0867

The average probability of survival at time interval l for patient no. d in group g can be estimated as

$$S_g(l) = \frac{1}{n_l^{[g]}} \sum_{d=1}^{n_l^{[g]}} \Pr_d^{[g]}(l), \quad g = 1, 2, l = 1, 2, \dots, \quad (16)$$

where $n_l^{[g]}$ is the total number of patients at time interval l in group g and $\Pr_d^{[g]}(l)$ is the survival function $\Pr(l)$ of patient no. d at time interval l in group g ; see, for example, Thomsen et al. [29].

The data are analyzed to discover which values of the covariates are likely to be of benefit. We compare the results obtained using smoothing spline, partial logistic, Cox’s proportional hazards, and piecewise exponential models [7, 8]. The results of fitting the various models are summarized in Table 3. It is clear from Table 3 that

- (i) all covariates for the smoothing spline model are strongly significant (in particular, Crowley and Hu [7] suggested a quadratic effects of age) and
- (ii) there is little difference between Cox’s proportional hazard model and the piecewise exponential model. It should be noted that binary covariates in the model remain linear.

As shown in Aitkin et al. [8, Figure 2], it is more appropriate to compare survivorship functions if the hazards are not proportional. One point of interest is a comparison of survival experience of transplanted and nontransplanted patients. Our proposal for comparing the survival function is to use the estimated survival function for only 41 heart transplanted patients who died to assess the efficacy of transplantation and the effects of covariates by means of modeling the change in hazard at transplantation by using (15) and (16). Our particular interest is the effect of waiting time on posttransplant survival according to several models. In Figure 4, two time periods are used (group 1: up to 20 days; group 2: longer than 20 days). Figure 4 shows a comparison of the estimated survival function. The estimated survival functions based on the smoothing spline suggest that patients with a short waiting time face a greater early risk than those who had a longer waiting time. However the estimated survival functions based on piecewise exponential models cannot reveal the difference between a short and long waiting times. Our method provides an alternative to Arjas’ [10] suggestion of comparing separate estimates of cumulative hazard based on the levels of the waiting time. Although Arjas [10] did not include waiting time as a covariate in Cox’s proportional hazard model because of nonproportionality issues, we used transplant status (i.e., waiting time), which is

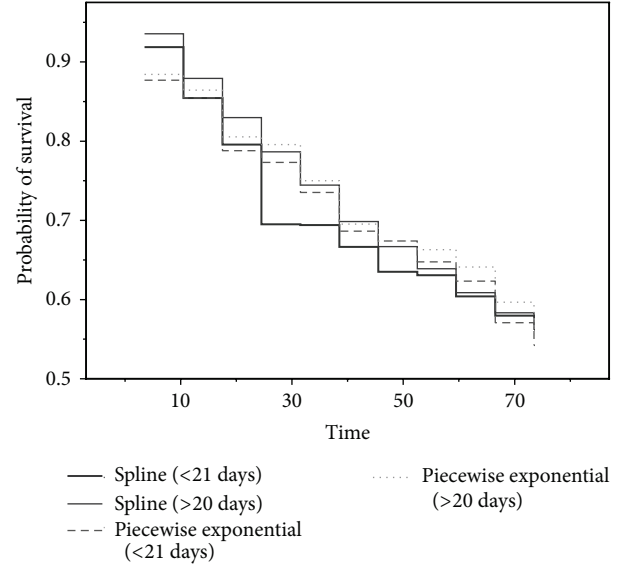


FIGURE 4: Survival function describing the effect of the waiting time for 41 heart transplanted patients who died.

strongly significant for the smoothing spline model according to the results shown in Table 3.

A fundamentally different type of analysis was suggested by Crowley and Hu [7] to investigate the effect of transplantation with a low mismatch score. They pointed out that transplantation may be beneficial for younger patients only based on regression coefficients for Cox’s proportional hazards model, but our conclusion can be derived by graphical analysis as well as significance testing of covariates. Defining a low mismatch score as less than or equal to one for all 29 heart transplanted patients [7], Figure 5 shows a graphical comparison of the estimated survival function for two groups, namely, the younger patients (less than 50 years old at acceptance) and older patients (greater than or equal to 50 at acceptance). From Figure 5, it is clear that older patients face a greater early risk than younger patients; see, for details, Crowley and Hu [7, Chapter 5] with respect to the cutpoints for low mismatch score as less than or equal to one and the younger patients as less than 50 years old. Kalbfleish and Prentice [30, Section 4.6.3] estimated the cutpoint for age, based on all 65 transplanted patients, as 46.2. Figure 6 shows a graphical comparison of the estimated survival function for two groups, namely, the younger patients (less than or equal 46 years old at acceptance) and older patients (greater than 46 at acceptance). As Kalbfleish and Prentice point out, transplantation is beneficial for younger patients.

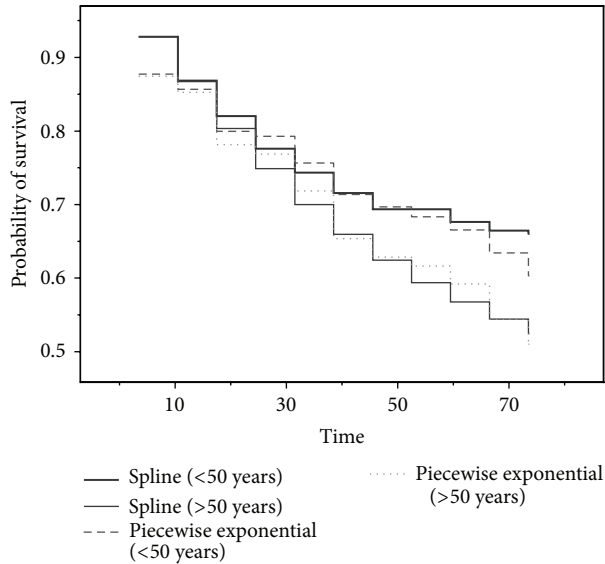


FIGURE 5: Survival function describing the effect of the age of transplantation for patients with a low mismatch score.

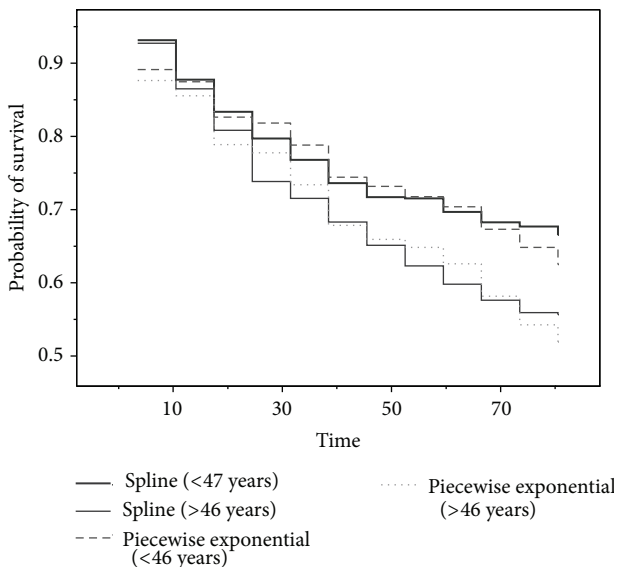


FIGURE 6: Survival function describing the effect of the age for all 65 transplanted patients.

4. Conclusion

We confined our attention to time-dependent covariates. Allowing covariates to vary over the duration of the study not only enabled us to study time-varying risk factors, but also provided a flexible way for modeling censored survival data using penalized smoothing splines. We illustrated the procedures using data of the Stanford Heart Transplant data.

By introducing the maximum likelihood principle into GAM,

- (i) we could visualize the spline effects of the midpoint of the time interval;

- (ii) the smoothing parameters could be selected by using variant ν -fold CV;
- (iii) the goodness-of-fit of GAM could be tested based on bootstrapping;
- (iv) the estimated average probabilities of survival enabled us to investigate the effect of transplantation with a low mismatch score for two groups, namely, the younger and older patients.

References

- [1] J. P. Klein and M. L. Moeschberger, *Survival Analysis*, Springer, New York, NY, USA, 2nd edition, 2003.
- [2] D. Collett, *Modelling Survival Data in Medical Research*, Chapman and Hall, London, UK, 1994.
- [3] P. A. Murtaugh, E. R. Dickson, G. M. van Dam et al., "Primary biliary cirrhosis: prediction of short-term survival based on repeated patient visits," *Hepatology*, vol. 20, no. 1 I, pp. 126–134, 1994.
- [4] E. Christensen, P. Schlichting, P. K. Andersen et al., "Updating prognosis and therapeutic effect evaluation in cirrhosis with Cox's multiple regression model for time-dependent variables," *Scandinavian Journal of Gastroenterology*, vol. 21, pp. 163–174, 1986.
- [5] E. Christensen, D. G. Altman, J. Neuberger et al., "Updating prognosis in primary biliary cirrhosis using a time-dependent Cox regression model," *Gastroenterology*, vol. 105, no. 6, pp. 1865–1876, 1993.
- [6] D. G. Altman and B. L. de Stavola, "Practical problems in fitting a proportional hazards model to data with updated measurements of the covariates," *Statistics in Medicine*, vol. 13, no. 4, pp. 301–341, 1994.
- [7] J. Crowley and M. Hu, "Covariance analysis of heart transplant survival data," *Journal of the American Statistical Association*, vol. 72, pp. 27–36, 1977.
- [8] M. Aitkin, N. Laird, and B. Francis, "A reanalysis of the Stanford Heart Transplant data," *Journal of the American Statistical Association*, vol. 78, pp. 264–292, 1983.
- [9] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, John Wiley, New York, NY, USA, 2nd edition, 2003.
- [10] E. Arjas, "A graphical method for assessing goodness of fit in Cox's proportional hazards model," *Journal of the American Statistical Association*, vol. 83, pp. 204–212, 1988.
- [11] E. Biganzoli, P. Boracchi, L. Mariani, and E. Marubini, "Feed forward neural networks for the analysis of censored survival data: a partial logistic approach," *Statistics in Medicine*, vol. 17, pp. 1169–1186, 1998.
- [12] E. Biganzoli, P. Boracchi, and E. Marubini, "A general framework for neural network models on censored survival data," *Neural Networks*, vol. 15, no. 2, pp. 209–218, 2002.
- [13] M. Tsujitani and M. Sakon, "Analysis of survival data having time-dependent covariates," *IEEE Transactions on Neural Networks*, vol. 20, no. 3, pp. 389–394, 2009.
- [14] M. Tsujitani, Y. Tanaka, and M. Sakon, "Survival data analysis with time-dependent covariates using generalized additive models," *Computational and Mathematical Methods in Medicine*, vol. 2012, Article ID 986176, 9 pages, 2012.
- [15] S. W. Lagakos, "The graphical evaluation of explanatory variables in proportional hazard regression models," *Biometrika*, vol. 68, no. 1, pp. 93–98, 1981.

- [16] D. R. Cox, "Partial likelihood," *Biometrika*, vol. 62, no. 2, pp. 269–276, 1975.
- [17] B. Efron, "Logistic regression, survival analysis, and Kaplan-Meier curve," *Journal of the American Statistical Association*, vol. 83, pp. 414–425, 1988.
- [18] T. J. Hastie and R. J. Tibshirani, *Generalized Additive Models*, Chapman and Hall, London, UK, 1990.
- [19] T. J. Hastie, R. J. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, New York, NY, USA, 2001.
- [20] P. McCullagh and J. A. Nelder, *Generalized Linear Models*, Chapman and Hall, London, UK, 2nd edition, 1989.
- [21] P. K. Andersen and R. D. Gill, "Cox's regression model for counting process: a large sample study," *Annals of Statistics*, vol. 10, pp. 1100–1120, 1982.
- [22] S. N. Wood, "Stable and efficient multiple smoothing parameter estimation for generalized additive models," *Journal of the American Statistical Association*, vol. 99, no. 467, pp. 673–686, 2004.
- [23] S. N. Wood, *Generalized Additive Models: An Introduction with R*, Chapman and Hall, London, UK, 2006.
- [24] S. N. Wood, "Fast stable direct fitting and smoothness selection for generalized additive models," *Journal of the Royal Statistical Society B*, vol. 70, no. 3, pp. 495–518, 2008.
- [25] P. Zhang, "Model selection via multifold cross validation," *Annals of Statistics*, vol. 21, pp. 299–313, 1993.
- [26] D. Collett, *Modelling Binary Data*, Chapman and Hall, London, UK, 2nd edition, 2003.
- [27] J. M. Landwehr, D. Pregibon, and A. C. Shoemaker, "Graphical methods for assessing logistic regression models," *Journal of the American Statistical Association*, vol. 79, pp. 61–71, 1984.
- [28] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman and Hall, New York, NY, USA, 1993.
- [29] B. L. Thomsen, N. Keiding, and D. G. Altman, "A note on the calculation of expected survival, illustrated by the survival of liver transplant patients," *Statistics in Medicine*, vol. 10, no. 5, pp. 733–738, 1991.
- [30] J. D. Kalbfleisch and R. L. Prentice, *The Statistical Analysis of Failure Time Data*, John Wiley, New York, NY, USA, 2nd edition, 2002.