Kauermann, Berger:

# A Smooth Test in Proportional Hazard Survival Models using Local Partial Likelihood Fitting

Sonderforschungsbereich 386, Paper 282 (2002)

Online unter: http://epub.ub.uni-muenchen.de/





# A Smooth Test in Proportional Hazard Survival Models using Local Partial Likelihood Fitting

Göran Kauermann — University of Glasgow

Ursula Berger* — MRC Glasgow

8th April 2002

**Abstract**

Proportional hazard models for survival data, even though popular and numerically handy, suffer from the restrictive assumption that covariate effects are constant over survival time. A number of tests have been proposed to check this assumption. This paper contributes to this area by employing local estimates allowing to fit hazard models with covariate effects smoothly varying with time. A formal test is derived to test the model with proportional hazards against the smooth general model as alternative. The test proves to possess omnibus power. Comparative simulations and two data examples accompany the presentation. Extensions are provided to multiple covariate settings, where the focus of interest is to decide which of the covariate effects vary with time.

KEYWORDS: Cox Model, Local Likelihood, Proportional Hazard Model, Smooth Tests

*Göran Kauermann, Department of Statistics & Robertson Centre, University of Glasgow, University Gardens, Glasgow G12 8QW, Scotland. Ursula Berger, Medical Research Council, Social and Public Health Science Unit, University of Glasgow, 4 Lillybank Gardens, Glasgow G12 8RZ, Scotland

# 1 Introduction

In the last decade a remarkable large number of publications considered model validation of parametric regression models using smoothing techniques. The methods are applicable for models where the influence of a continuous covariate, e.g. time, is modelled in a parametric fashion. The basic idea behind smooth tests is to go beyond the parametric framework by including non-parametric but smooth components in the model. Inference is then drawn by comparing the fit of both models by appropriate means. The classical and most simple example is the linear normal response regression model where the parametric mean structure can be checked and tested against a general but smooth model. References for smooth tests in regression models include among others e.g. le Cessie & van Houwelingen (1991), Müller (1992), Härdle & Mammen (1993), Aerts, Claeskens & Hart (1999) and Härdle & Kneip (1999) and citations given in those papers. A general overview of existing methods is found in Hart (1997). Recently, Bowman & Young (1996) and Kauermann & Tutz (2001) extended the testing problem to investigate whether the effect of a factorial covariate is modified smoothly by a continuous covariate, e.g. time. This is a typical problem occurring in survival analysis where factorial covariate effects may vary over survival time.

A popular model for survival data is the proportional hazard (PH) model due to Cox (1972). The PH model is powerful and numerically handy, but in it's standard form it assumes the proportionality of the hazard functions, i.e. covariate effects are supposed to be constant over survival time. In order to investigate and test this assumption several routines have been suggested. Parametric extensions by including dynamic time-covariate effects with corresponding tests based on the estimated co-

efficients have been suggested e.g. in Cox (1972) and Grambsch & Therneau (1994). Such tests require the pre-specification of a suspected departure from proportionality in a functional form. The demand for pre-specification of time-covariate interactions makes tests of this kind less flexible for detecting complex departures from proportionality. Smooth estimation of covariate effects in the PH model using smoothing splines has been suggested e.g. by O'Sullivan (1988) or Hastie & Tibshirani (1990a) in an exploratory fashion. Hastie & Tibshirani thereby propose to make use of the deviance to compare or 'screen' models with time-varying covariate effects, even though they already mention that asymptotic theory which justifies this approach does not exist. Gray (1994) further develops penalised B-spline fitting towards testing. He makes use of low dimensional smoothers which allows to apply asymptotic results for quadratic forms to calculate the $p$-value. A similar idea is pursued in Hess (1994) or Abrahamowicz, MacKenzie & Esdaile (1996) who use regression splines respectively polynomial estimation. Inference is based on the estimated coefficients and the resulting variance matrix. Verweij & van Houwelingen (1995) use penalised regression by letting each event time-point have its own parameter. A first local approach for testing the PH assumption is suggested in Moreau, O'Quigley & Mesbah (1985) and O'Quigley & Pessione (1989) who fit local piecewise constant parameters capturing possible time-covariate interaction. Widely speaking, this idea can be seen as an ancestor of the local smoothing approach pursued in this paper. Local estimation of smooth effect functions in PH models is discussed in Fan, Gijbels & King (1997) and Cai & Sun (2002). The focus there is mainly on estimation while we here concentrate on testing. For a general overview for estimation and tests in proportional hazard models we also refer to Lin & Wei (1991) or Sasieni (1999).

In this paper we present a smooth test for the proportional hazard assumption based on local partial likelihood estimates. It can be seen as an extension of (Kauermann & Tutz, 2001) applied to survival data. As test statistic we consider the log partial likelihood ratio comparing a smooth fit and a parametric fit. The approach is flexible and allows to uncover even complex departures from the proportionality of the hazards. Smooth test statistics based on local estimates generally do not lead to simple reference distributions (see also Hastie & Tibshirani, 1990b, page 156). This is in particular since smoothers respectively smoothing matrices are not idempotent operators like projection matrices. We therefore propose to bootstrap the test statistic to obtain the reference distribution under the hypothetical model. In order to limit the computational burden this in turn demands to develop a numerically simple and fast smooth estimate. For this reason we suggest to make use of *local one step* estimates which prove to be numerically handy and sufficiently exact for testing purposes at the same time. The bootstrap itself is pursued following the procedure suggested in Davison & Hinkley (1997).

The paper is organised as follows. In Section 2 we introduce local partial likelihood estimation and derive a smooth test based on the local partial likelihood ratio statistic. Simulations and a small example demonstrate the applicability and performance of the routine. Section 3 extends the test to multi-variable situations where covariate effects are tested separately on time interaction. Again, a simulation study and a data example support our developments. A discussion concludes the paper and technical details are provided in the Appendix.

# 2 Smooth Tests in Survival Models

## 2.1 Dynamic Cox Models

The Cox model defines the proportional hazard for the $j$-th unit as

$$\lambda(t|X_j) \quad = \quad \lambda_0(t)\exp\{X_j\beta\}, \tag{1}$$

where $X_j$ is a set of covariates or risk factors and $\lambda_0(t)$ is an unspecified baseline hazard. Model (1) implies proportionality of the hazards, i.e. the ratio of two hazards from units $j$ and $k$ is constant over time with value $\exp\{(X_j - X_k)\beta\}$. It is obvious that this restriction is often too stringent and should be relaxed by allowing the covariate effects to vary with time. This leads to the dynamic Cox model

$$\lambda(t|X_j) \quad = \quad \lambda_0(t)\exp\{X_j\beta(t)\} \tag{2}$$

where $\beta(t)$ is a vector of smooth but unknown functions in $t$. Models of the kind (2) are generally introduced as varying coefficient models by Hastie & Tibshirani (1993). The shape of $\beta(t)$ thereby mirrors the interaction between time and the covariates. If $\beta(t)$ is constant, i.e. $\beta(t) \equiv \beta$, model (2) simplifies to (1). The objective in the following is now on testing the proportional hazard model (1) against its smooth extension (2).

## 2.2 Local Partial Likelihood Estimation

Parameter vector $\beta$ in (1) can be estimated by partial likelihood. This in particular circumvents the estimation of the baseline hazard. Let $T_j$ denote the survival time of the $j$ individual or observational units and let $C_j$ be the corresponding right censored time, $j = 1, \ldots, N$. We observe $Y_j = \min(T_j, C_j)$ and define the censoring

indicator $\delta_j = 1$ if $T_j < C_j$ and $\delta_j = 0$ otherwise. Let $t_1, \ldots, t_n$ denote the observed time-points of failure. With $\mathcal{D}_i$ we define the index set of units failing at time-point $t_i$, i.e. $\mathcal{D}_i = \{j : Y_j = t_i \text{ and } \delta_j = 1\}$ and $\mathcal{R}_i$ denotes the index set of units at risk at time-point $t_i$, i.e. $\mathcal{R}_i = \{j : Y_j \geq t_i\}$. The partial log likelihood is thereby defined as $\mathbf{l}(\beta) = \sum l_i(\beta|\mathcal{R}_i)$ with

$$l_i(\beta|\mathcal{R}_i) \quad = \quad \left( \sum_{j \in \mathcal{D}_i} X_j \beta \right) - |\mathcal{D}_i| \log \left\{ \sum_{j \in \mathcal{R}_i} \exp(X_j \beta) \right\} \tag{3}$$

with $|\mathcal{D}_i|$ as cardinality of the set $\mathcal{D}_i$. The form (3) is by itself an approximation which is due to Breslow (1974). In case of many ties, i.e. if $|\mathcal{D}_i|$ is large, (3) is known to produce biased estimates and should be replaced by more accurate approximations (see e.g. Hertz-Picciotto & Rockhill, 1997). For notational and computational simplicity we here stick to (3) however.

Let $s_i(\beta|\mathcal{R}_i)$ denote the score contribution

$$s_i(\beta|\mathcal{R}_i) \quad = \quad \frac{\partial l_i(\beta|\mathcal{R}_i)}{\partial \beta} = \widetilde{X}_i^T - |\mathcal{D}_i| \sum_{j \in \mathcal{R}_i} X_j^T \pi(j|\mathcal{R}_i, \beta) \tag{4}$$

where $\widetilde{X_i} = \sum_{j \in \mathcal{D}_i} X_j$ and $\pi(j|\mathcal{R}_i, \beta) = \exp(X_j\beta) / \sum_{k \in \mathcal{R}_i} \exp(X_k \beta)$. The partial likelihood estimate for $\beta$ is now obtained by solving the estimating equation

$$0 \quad = \quad \sum_{i=1}^{n} s_i(\widehat{\beta}_0|\mathcal{R}_i) \tag{5}$$

where subscript 0 here and sequel is used to refer to estimates for constant effects of the proportional hazard model (1). In order to smoothly fit dynamic effects in model (2) we solve (5) in a locally weighted manner. Consider therefore some kernel weights $w_{li} = K\{(t_l - t_i)/h\}$ with $K(\cdot)$ as unimodal kernel function and $h$ as bandwidth or smoothing parameter. Incorporating these weights in (5) provides

the *local partial score equation*

$$0 \quad = \quad \sum_{i=1}^{n} w_{li} s_i(\widehat{\beta}_l | \mathcal{R}_i) \tag{6}$$

which yields with $\widehat{\beta}_l$ an estimate for the dynamic effect $\beta(t_l)$. Solving (6) for all failure time-points provides the smooth estimate $\widehat{\beta}(t)$. Alternatively one may solve (6) for a grid of time points accompanied by subsequent interpolation. Asymptotic properties of estimate $\widehat{\beta}_l$ are derived in Cai & Sun (2002) (see also Fan, Gijbels & King, 1997). Some further details are provided in Appendix B. Similar to standard smoothing, the variance of the estimate can be approximated, under appropriate assumptions, by the weighted Fisher type matrix, i.e.

$$\text{var}(\widehat{\beta}_l) \quad = \quad c \left\{ -\sum_{i=1}^{n} w_{li} \frac{\partial s_i(\beta_i | \mathcal{R}_i)}{\partial \beta^T} \right\}^{-1} \tag{7}$$

with $c = \int K^2(t)dt$. Since constant $c$ depends on the kernel only, it's calculation is straight forward, e.g. for a normal kernel it is $c = 1/\sqrt{2}$.

The bandwidth $h$ in weight $w_{li}$ in (6) steers the amount of smoothing. It is noteworthy that smooth and parametric estimation are nested in that for $h \to \infty$ one gets $\widehat{\beta}_l \to \widehat{\beta}_0$. In principle, bandwidth $h$ can be selected data driven by cross validation or minimisation of the Akaike criterion or other tools. For testing purposes, however, the choice of an optimal bandwidth is of secondary focus and, as will be seen in simulations, the performance of the proposed test is only weakly sensible to the specific choice of $h$.

## 2.3 Smooth Testing on Dynamic Effects

The objective is now to compare the parametric fit $\widehat{\beta}_0$ with the smooth fit $\widehat{\beta}(t)$ in order to test the validity of the proportional hazard model (1). This means we

investigates whether the dynamic model (2) improves the fit significantly compared to (1). To do so we employ the partial log likelihood ratio statistic

$$\Lambda(h) \quad = \quad 2\sum_{i=1}^{n}\{l_i(\widehat{\beta}_i|\mathcal{R}_i) - l_i(\widehat{\beta}_0|\mathcal{R}_i)\} \tag{8}$$

where $h$ denotes the dependence on the bandwidth. It is shown in Appendix B that $\Lambda(h)$ has positive expectation which decomposes essentially to

$$E\{\Lambda(h)\} \quad = \quad E_{H_0}\{\Lambda(h)\} + \delta_h^2,$$

with $E_{H_0}(\cdot)$ denoting the expectation under the $H_0$ model (1) and $\delta_h^2$ as shift or non-centrality parameter. If model $H_0$ holds non-centrality vanishes, i.e. $\delta_h^2 \equiv 0$. Hence, in order to test model $H_0$ one has to assess the size of $\Lambda(h)$ with respect to it's reference distribution under $H_0$. As mentioned beforehand such a reference distribution is not directly available in an analytic form since asymptotic $\mathcal{X}^2$ or other normal approximations do not hold (see also Hastie & Tibshirani, 1990b, page 156). This is in particular in contrast to low dimensional and parametric smoothers where asymptotic distributions can be obtained via standard asymptotics and quadratic forms (see e.g. Gray, 1994). To circumvent the problem we make use of Bootstrapping to obtain the distribution of $\Lambda(h)$ under $H_0$. Bootstrapping right censored survival data is described in detail in Davison & Hinkley (1997) chapters 3.5 and 7.3. Some details are provided in Appendix A.

The use of bootstrapping demands to reduce the computational effort of estimation, since in each bootstrap iteration a smooth alternative model has to be fitted. We therefore replace the local estimates in the local score function (6) by a numerically handy one-step estimate. This means, instead of solving (6) exactly, we use

the local one-step Fisher scoring estimate

$$\widehat{\beta}_l^{(1)} \quad = \quad \widehat{\beta}_0 - \left\{ \sum_{i=1}^{n} w_{li} \frac{\partial s_i(\widehat{\beta}_0 | \mathcal{R}_i)}{\partial \beta^T} \right\}^{-1} \sum_{i=1}^{n} w_{li} s_i(\widehat{\beta}_0 | \mathcal{R}_i) \tag{9}$$

for the alternative model with $\widehat{\beta}_0$ as parametric estimate in the $H_0$ model (1). If $H_0$ holds, it is easy to see that $\widehat{\beta}_l^{(1)}$ is consistent, since $\widehat{\beta}_0$ is consistent. If on the other hand $H_1$ holds, $\widehat{\beta}_l^{(1)}$ uncovers the dynamic structure which is not coped for in $H_0$. We refer to Appendix B for more details. Apparently, $\widehat{\beta}_l^{(1)}$ is numerically simple and fast so that using it in the bootstrapping step reduces the computational burden noticeably. Additionally, estimates (9) do not necessarily have to be calculated at all observed failure time-points, but instead they can be calculated on a grid of time-points with subsequent interpolation. This again reduces the computational effort and makes the procedure easy and numerically feasible.

## 2.4 Simulation and Example

*Simulation*

We run a simulation study to assess the performance of our proposed test. We take $X$ as single binary covariate with orthogonal design, i.e. we simulate 100 survival times with $X = 1$ and $X = 0$, respectively. The survival time is generated from a logistic setting with time $t$ taking discrete values $1, 2, \ldots$ and

$$\text{logit } P(T = t | T \geq t, x) \quad = \quad \lambda_0(t) + X\beta(t)$$

with constant baseline hazard $\lambda_0(t) = \text{logit}^{-1}(-4)$ (for asymptotic equivalence of the Logit model and the Cox model see Thompson 1977). Censoring is simulated to be independent of $X$ and $t$ with probability 0.005 at each time point, yielding a censoring rate of about 25% (for the baseline group $X = 0$). Exemplary we show in

Figure 1 (right two plots) the estimated survivor and censoring function (based on a fitted Cox model) resulting from one simulation.

In the first simulation the effect $\beta(t)$ is kept time constant (model $H_0$) for assessing the consistency of the test. To study the power of the test, seven different dynamic settings are considered, where $\beta(t)$ is taken as linear, quadratic, cosinus or discontinuous step function, as listed in Table 1 and visualised in Figure 1 (left plot).

For each of these setting we generate 250 replicates. The smooth tests is determined based on 100 bootstraps with bandwidth $h = 40$. Figure 2 shows the distribution functions of the $p$-values for the different settings where the identity-line is given as reference. Clearly, the test appears to be consistent and provides an omnibus type power against the different alternatives. In Table 1 we also list the simulated rejection probabilities based on the significance levels 0.05 % and 0.1 %.

For comparison, we apply a number of other PH-tests in four of the above simulation settings (constant, flat linear, flat quadratic, weak cosinus). This includes the approach suggested in (Cox, 1972) where the effect of a pre-specified time-covariate interaction, here a linear one, is tested. Related to this is the score based test proposed by Grambsch & Therneau (1994) (GT) which is also based on testing a pre-specified dynamic structure. Not surprisingly due to the similarity of their construction both provide nearly the same results as seen from the simulated distribution function of the $p$-value shown in Figure 3. Secondly we use the piecewise approach suggested in Moreau, O'Quigley & Mesbah (1985) or O'Quigley & Pessione (1989). This is based on fitting constant effects over pre-specified segments of

time, where we here partition time at four equidistant knots. The piecewise test is rather conservative and despite a rather high partitioning clearly not flexible enough to capture complex departures of the PH-assumption. Finally the spline based test suggested in Hess (1994) is applied. This test makes use of cubic regression splines and can be seen as a smooth and flexible extension of the piecewise approach. However, the test requires to choose the number and location of knots, where we here take 3 and 5 knots (spline(3) and spline (5)). The distribution of the p-values for this test reveals some dependency on the number of knots. Moreover, the test becomes too liberal if more knots are included.

Generally, it appears that the local approach suggested in this paper performs promising and in fact outperforms some of the considered competitors in the different simulation settings. In the quadratic situation it is beaten by the spline approaches, which however show a too liberal performance in the $H_0$ scenario. A general superiority of our test can of course not be drawn from this limited simulation study, we think however there is enough evidence that the test shows a rather satisfactory behavior.

Finally, we investigate the robustness of our test with respect to the chosen bandwidth $h$. We therefore repeat the test for the constant, the linear and the quadratic setting using the different bandwidths $h = 30, 60$ and 120. The trajectories of the $p$-values for some simulations from the linear setting are shown in Figure 4. Apparently the variation of the $p$-value for different bandwidths is small. Table 2 lists the matching of conclusions drawn from the test for different bandwidths. The different 2×2 tables read as follows. For instance in 93% of the simulations from model $H_0$ both tests based on $h = 30$ and $h = 60$ accept the hypotheses

while it is rejected by both tests in 5% of the case. Clearly there is evidence for matching results, even though for the quadratic setting the power is clearly reduced for smaller bandwidths. Generally, the simulations show that the test procedure is rather insensible to different values of the bandwidth. The same pattern is also observed in the following example.

*Example*

To exemplify the routine we apply the method to the gastric cancer data used and listed in Hess (1994). The data consist of 90 patients with non-resectable gastric cancer who either receive chemotherapy (control group, n=45) or a combination of chemotherapy and radiation (n=45). Hess reports a violation of the PH-assumption for the therapy effect with a $p$-value of 0.015 for his test when using his spline based test. Our test based on local partial likelihood estimation confirms this finding. The $p$-values resulting from 500 bootstraps for different bandwidths are 0.002 ($h = 250$), 0.004 ($h = 500$) and 0.010 ($h = 1500$) and they clearly indicate time-treatment interaction. Similar to the simulation study above we observe that the $p$-values are roughly the same for different bandwidths. Figure 5 shows the resulting local partial likelihood estimates $\widehat{\beta}(t)$ for the treatment effect for the three different bandwidths. The parametric fit based on the PH-assumption is given as reference line. Even though the fit for bandwidth $h = 1500$ is hardly distinguishable from the parametric fit, the difference proves to be significant. This again demonstrates the robustness of the testing approach against the bandwidth used. The example also demonstrates that while estimation depends on the chosen bandwidth (and hence requires a data driven choice of the bandwidths) the results of the test based on local partial likelihood is only weakly dependent on the selection of the bandwidth.

# 3 Componentwise Smooth Tests

## 3.1 Semiparametric Models

In a multivariable situation the focus of interest is not only to investigate whether the proportional hazard model (1) is appropriate or not, but also which of the covariate effects, if any, varies with time. This problem extends the previous setting in that we now allow some of the covariates to have dynamic effects while others have time constant effects. The alternative model then becomes

$$\lambda_j(t|X_j) \quad = \quad \lambda_0(t)\exp\{X_{1j}\beta_1(t) + X_{2j}\beta_2\} \tag{10}$$

where matrix $X_{1j}$ includes those covariates which effects are allowed to vary over time and $X_{2j}$ comprises those covariates which have time constant effects.

Fitting the semiparametric model (10) can be done by backfitting, i.e. fitting successively the fixed effect $\beta_2$ by partial likelihood and the varying effect $\beta_1(t)$ by local likelihood, treating the other parameters respectively as given. Backfitting however is time consuming since iteratively a fixed effect model and a varying coefficient model has to be fitted, each containing a given offset resulting from the estimate of the previous backfitting loop. This basically excludes the procedure to be applied in our testing context since the computational burden would not be acceptable. For testing purposes however a highly accurate fit is not necessary since the focus is on model validation and not on efficient point estimation. Therefore, as above, we suggest to make use of a one step backfitting estimate which reduces the computational effort significantly by still maintaining a satisfactory performance of the smooth test.

Let $\widehat{\beta}_0 = (\widehat{\beta}_{10}^T, \widehat{\beta}_{20}^T)^T$ be the parametric fit resulting from estimating the proportional hazard model (1) with $X = \{X_1^T, X_2^T\}^T$. We fix $\widehat{\beta}_{20}$ and consider the one step estimate of the solution to the local backfitting estimating equation for $\beta_{1l} = \beta_1(t_l)$

$$0 \quad = \quad \sum_{i=1}^{n} w_{li} s_{1i}(\widehat{\beta}_{1l}; \widehat{\beta}_{20} | \mathcal{R}_i) \tag{11}$$

with $s_{1i}(\beta_1; \widehat{\beta}_{20} | \mathcal{R}_i) = \sum_{j \in \mathcal{D}_i} X_{1i}^T - |\mathcal{D}_i| \sum_{j \in \mathcal{R}_i} X_{1j}^T \pi\{j | \mathcal{R}_i, (\beta_1^T, \widehat{\beta}_{20}^T)^T\}$ and $\pi()$ as defined subsequent to (4). Similar to the previous section this leads to the one step estimate

$$\widehat{\beta}_{1l}^{(1)} \quad = \quad \widehat{\beta}_{10} - \left\{ \sum_{i=1}^{n} w_{li} \frac{\partial s_{1i}(\widehat{\beta}_{10}; \widehat{\beta}_{20} | \mathcal{R}_i)}{\partial \beta_1^T} \right\}^{-1} \sum_{i=1}^{n} w_{li} s_{1i}(\widehat{\beta}_{10}; \widehat{\beta}_{20} | \mathcal{R}_i).$$

The estimate is numerically simple and easy applicable to be used in bootstrapping. Inserting the resulting one step estimates in the likelihood ratio yields

$$\Lambda_1(h) = 2 \left[ \sum_{i=1}^{n} l_i\{(\widehat{\beta}_{1i}^T, \widehat{\beta}_{20}^T) | \mathcal{R}_i\} - l_i(\widehat{\beta}_0 | \mathcal{R}_i) \right] \tag{12}$$

as test statistic for assessing the PH-assumption for $X_1$. For calculation of the $p$-value we simulate $\Lambda_1(h)$ under $H_0$ by bootstrapping as suggested in the previous section.

## 3.2 Example and Simulation

*Simulation*

We simulate survival data as in the previous section using a logistic setting but now with hazard function

$$\text{logit } P(T = t | T \geq t, x) \quad = \quad \lambda_0(t) + x_1 \beta_1(t) + x_2 \beta_2.$$

The baseline is again constant as in the simulation in the previous section and we set $\beta_2 = 0$. The covariates $x_1$ and $x_2$ are independent with balanced design for $x_1$ and

$x_2$ randomly drawn with $P(x_2 = 1) = 0.4$. We simulate 250 datasets each with 200 individuals with (a): $\beta_1(t) \equiv 0$ to investigate consistency and (b): $\beta(t)$ as (weak) quadratic function (see Table 1) to explore the power. Figure 6 gives the lower part of the distribution function of the p-values for the global PH-test $\Lambda(h)$. Moreover we show the distribution functions of the p-values when testing the PH- assumption for single components using $\Lambda_1(h)$ as given in (12), and of $\Lambda_2(h)$ defined as in (12) but with indices 1 and 2 exchanged. Clearly the test is consistent in (a) and shows power in (b) by rejecting the PH model. Moreover, the component-wise tests allow to explore which of the covariates shows significant interaction with time.

*Example*

In a final example we illustrate the use of the proposed test procedure to find those covariates which show distinct interaction with time. The data come from a breast cancer study on the clinical relevance of two tumour-biological factors, the urokonase-type Plasminogen Activator uPA and its type-one inhibitor PAI-1, both coded as binary indicators taking value 1 if the level of measurement exceeds a given cut-point (for details see Harbeck et al., 1999, 2001). The two factors uPA and PAI-1 were determined in primary tumour tissue extracts from 316 breast cancer patients who showed no evidence of distant metastases. The effect of uPA and PAI-1 on disease free survival after surgery is investigated in a multi-variable analysis where in addition a number of classical prognostic factors are taken into account. These include the nodal state at surgery (LYPO) as binary indicator taking value 1 if the axillary lymph nodes showed tumour cell involvement, the positive or negative hormone receptor state (HORMO) incorporating estrogen and progesteron receptor expression, and high or low rating of tumour grading (GRADI) reflecting the

proliferation of the tumour.

Figure 7 shows the local-likelihood estimates of the time-varying effects (using a bandwidths of $h = 40$) together with the 95% confidence intervals based on (7). In addition, the estimated effects in the PH model are given (horizontal lines). The graphical presentation of the effects reveals some time-variation in the effects of PAI-1 and hormo and possibly uPA. A formal analysis to test on the dynamic structure of these effects is carried out using the proposed test. The p-values for the global test are listed in Table 3 for different bandwidths. They are all $< 0.01$ and unanimously point out the inadmissibility of the PH- assumption. For a detailed analysis exploring which of the factors interact with time, component-wise $p$-values are determined . This is done for each of the five factors again using different bandwidths. The resulting p-values are apparently homogeneous over the different bandwidths and the test rejects constant effects for the hormone receptor state and PAI-1 while other factors do not show any significant interaction with time.

## 4 Discussion

In this paper we proposed a test on proportional hazards based on local partial likelihood fitting. The $p$-value is calculated by bootstrapping where the computational burden is limited by one step estimation. The promising feature of the proposed test lies in its omnibus character yielding power against a wide range of alternatives.

The testing principle can be extended to more general goodness of fit tests to also investigate parametric time-covariate interaction. For example, if the linear interaction model $\lambda(t|x) = \lambda_0(t) \exp(x\beta_0 + x\ t\ \beta_t)$ is considered to be tested it can

be estimated locally by local linear partial likelihood estimation. The difference in the fit can again be assessed by bootstrapping.

# A Bootstrapping Survival Data

The following bootstrap procedure is essentially as suggested in Davison & Hinkley (1997). It is numerically implemented in the Splus procedure `censboot()` available from the server `http://statwww.epfl.ch/davison/BMA/`. The basic idea is to pursue a conditional bootstrap, conditional on the pattern of censorship. Let $1 - F_0(t|x;\beta_0) = P(T > t|x,\beta_0)$ be the survivor function of survival time $T$ based on the parametric hypothetical model (1). With $1 - \widehat{F}_0(t|x) = \{1 - \widehat{F}_0(t)\}^{\exp(x\widehat{\beta}_0)}$ we denote a corresponding estimate, where $\widehat{\beta}_0$ is the partial likelihood estimate and $\widehat{F}_0(t)$ is an estimate for the baseline distribution function obtained e.g. using the Breslow estimate (see e.g. Breslow, 1972 and Breslow, 1974). Moreover, let $\widehat{G}(t)$ be an estimate for the distribution function of the censoring time obtained by

$$1 - \widehat{G}(t) \;\; = \;\; \prod_{j \,:\, y_j \le t} \left\{ \frac{N-j}{N+1-j} \right\}^{1-\delta_j} .$$

In principle, $G(\cdot)$ can also be estimated model based (see Davison & Hinkley, 1997, page 351). Assuming random censoring, the distribution function of the observed survival time results by $H(t) = F_0(t)G(t)$ which is estimated by $\widehat{F}_0(\cdot)$ and $\widehat{G}(\cdot)$. Bootstrap replicates are then generated as follows.

1. Generate $T_1^*, \ldots, T_N^*$ i.i.d. from the parametric model, i.e. from the distribution function $\widehat{F}_0(\cdot)$.

2. For $\delta_j = 0$ set $C_j^* = Y_j$ and for $\delta_j = 1$ generate the censoring variable $C_j^*$

by drawing from the conditional distribution function $\widehat{G}(c|c > Y_j) = \{\widehat{G}(c) - \widehat{G}(Y_j)\}/\{1 - \widehat{G}(Y_j)\}$.

3. Define the observed bootstrapped survival time by $Y_j^* = \min(T_j^*, C_j^*)$.

4. Use the resulting bootstrap sample $Y_j^*, X_j$, $j = 1, \ldots, N$ to calculate the bootstrap likelihood ratio statistic $\Lambda^*(h)$.

Repeating the above steps $B$ times provides the bootstrap likelihood ratios $\Lambda^{*b}(h)$, $b = 1, \ldots, B$. The bootstrap $p$-value is then obtained by $\sum_{b=1}^{B} 1\{\Lambda^{*(b)}(h) > \widehat{\Lambda}(h)\}/B$ with $1\{\}$ as indicator function.

# B Technical Details

*Local Partial Likelihood Estimates*

We give a short sketch of properties of the local partial likelihood estimate. A formal theoretical derivation including necessary assumptions can be found in Cai & Sun (2002) (see also Fan, Gijbels & King, 1997). Assume that $\lambda_j(t) = \lambda_0(t) \exp\{X_j \beta(t)\}$. We decompose the varying coefficient to $\beta(t) = \beta + \xi(t)$ with $\beta$ fulfilling

$$0 = \sum_{i=1}^{n} E\{s_i(\beta|\mathcal{R}_i)\}. \tag{13}$$

This means essentially that $\beta$ is the best constant approximation of $\beta(t)$ in terms of the partial likelihood. Clearly, if model (1) holds $\xi(t) \equiv 0$. Hence, when testing (1) against (2) the component $\xi(t)$ is playing an important role for the power of the test. We abbreviate $\beta_i = \beta + \xi_i$ with $\xi_i = \xi(t_i)$ and write $s_i$ for $s_i(\beta_i|\mathcal{R}_i)$, i.e. we drop parameter arguments if the likelihood terms are calculated at the true parameter

values. The second order derivative is denoted by $\nabla s_i(\beta_i|\mathcal{R}_i)$, respectively $\nabla s_i$, where

$$\nabla s_i \;=\; -\sum_{j\in\mathcal{R}_i} X_j^T X_j \pi(j|\mathcal{R}_i,\beta_i) + \left\{\sum_{j\in\mathcal{R}_i} X_j^T \pi(j|\mathcal{R}_i,\beta_i)\right\}\left\{\sum_{j\in\mathcal{R}_i} X_j \pi(j|\mathcal{R}_i,\beta_i)\right\}.$$

Note that $\nabla s_i$ does not depend on $\mathcal{D}_i$. Under appropriate smoothness assumptions, expansion of the local partial score function(6) provides

$$\begin{aligned} 0 &= \sum_{i=1}^{n} w_{li} s_i(\widehat{\beta}_l|\mathcal{R}_i) \\ &= \sum_{i=1}^{n} w_{li}\left[s_i + \nabla s_i(\widehat{\beta}_l - \beta_i) + O\{(\widehat{\beta}_l - \beta_i)^2\}\right]. \end{aligned} \tag{14}$$

Replacing $(\widehat{\beta}_l - \beta_i)$ in (14) by $\{(\widehat{\beta}_l - \beta_l) + (\beta_l - \beta_i)\}$ allows to solve (14) for $\widehat{\beta}_l - \beta_l$ which yields

$$\widehat{\beta}_l - \beta_l \;=\; \{-\sum_{i=1}^{n} w_{li}\nabla s_i\}^{-1} \sum_{i=1}^{n} w_{li}\left[s_i + b_{li} + O\{(\widehat{\beta}_l - \beta_i)^2\}\right] \tag{15}$$

with bias component $b_{li} = \nabla s_i(\xi_l - \xi_i)$. Under $H_0$ the bias component vanishes, i.e. $b_{li} \equiv 0$. Assuming $\widehat{\beta}_l$ to be consistent it is easily seen from (15) that the asymptotic variance of $\widehat{\beta}_l$ equals

$$\begin{aligned} \mathrm{var}(\widehat{\beta}_l) &= \{\sum_{i=1}^{n} w_{li}\nabla s_i\}^{-1}\{\sum_{i=1}^{n} w_{li}^2 E(s_i s_i^T)\}\{\sum_{i=1}^{n} w_{li}\nabla s_i\}^{-1} + \ldots \\ &\approx c\{\sum_{i=1}^{n} w_{li}\nabla s_i\}^{-1} \end{aligned}$$

with $c \approx \sum_{i=1}^{n} w_{li}^2 / \sum_{i=1}^{n} w_{li} \approx \int K^2(z)dz$.

Using the same notation, we can also expand (5) in a similar fashion as above yielding

$$\widehat{\beta}_0 - \beta_l \;=\; \{-\sum_{i=1}^{n} \nabla s_i\}^{-1} \sum_{i=1}^{n}\left[s_i + b_{li} + O\{(\widehat{\beta} - \beta_i)^2\}\right]. \tag{16}$$

In (16), the bias component thereby decomposes asymptotically to

$$\{-\sum_{i=1}^{n}\nabla s_i\}^{-1}\sum_{i=1}^{n}b_{li} \;=\; \xi_l - \{-\sum_{i=1}^{n}\nabla s_i\}^{-1}\sum_{i=1}^{n}\nabla s_i\xi_i \;\approx \xi_l. \tag{17}$$

The latter simplification results from the definition of $\xi$ given through (13) by simple first order expansion of (13) about the true parameters leads to (17).

*Partial Likelihood Ratio*

Formulae (15) and (16) are now used to expand the likelihood ratio. Let $p$ be the dimension of $\beta$. We define the $(np)\times(np)$ dimensional smoothing type matrix $\boldsymbol{W}(h)$ which is build from $p$ dimensional block matrices

$$\boldsymbol{W}_{(li)} \;=\; w_{li}\left\{-\sum_{j=1}^{n}w_{lj}\nabla s_j\right\} \tag{18}$$

with $l = 1,\ldots,n$ and weights $w_{li} = K\{(t_l - t_i)/h\}$ depending on the bandwidth $h$. Making use of this notation estimate (15) can in first order approximation be written in matrix form as

$$\begin{aligned}\widehat{\beta}_l - \beta_l &= \sum_{i=1}^{n}\boldsymbol{W}_{(li)}(h)\{s_i + \nabla s_i(\xi_l - \xi_i)\} + \ldots \\ &= \boldsymbol{W}_{(l)}(h)\boldsymbol{s} + \boldsymbol{b}_l(h)\end{aligned} \tag{19}$$

where $\boldsymbol{W}_{(l)} = (\boldsymbol{W}_{(l1)},\ldots,\boldsymbol{W}_{(ln)})$ and $\boldsymbol{s} = (s_1^T,\ldots,s_n^T)^T$. Similarly, the bias component in (19) is available in matrix form by

$$\boldsymbol{b}_l(h) \;=\; \boldsymbol{W}_{(l)}(h)\mathrm{diag}_n(\nabla s_i)\boldsymbol{\xi} - \boldsymbol{\xi}$$

where $\boldsymbol{\xi} = (\xi_1^T,\ldots,\xi_n^T)^T$ and $\mathrm{diag}_n(\nabla s_i)$ denotes the block diagonal matrix having $\nabla s_i$, $i = 1,\ldots,n$ on its diagonal. In the same fashion we can obtain a first order expansion for the parametric estimate in model $H_0$ simply by setting $h \to \infty$. This means we obtain from (16)

$$\widehat{\beta}_0 - \beta_l \;=\; \boldsymbol{W}_{(li)}(\infty)s_i + \boldsymbol{b}_l(\infty)$$

with $\boldsymbol{W}_{li}(\infty) = \lim_{h\to\infty}\boldsymbol{W}_{li}(h) = \{-\sum_{j=1}^{n}\nabla s_j\}^{-1}$ and $\boldsymbol{b}_l(\infty) = \lim_{h\to\infty}\mathbf{b}_l(h)$. Note that $\boldsymbol{W}(h)^T\mathrm{diag}_n(\nabla s_i)\boldsymbol{W}(h) = -\boldsymbol{W}(h)$ which follows by simple calculation and is needed subsequently. The likelihood ratio can now be expanded by standard Taylor series, which yields by simple matrix algebra

$$\begin{aligned}\Lambda(h) &\approx \boldsymbol{s}^T\boldsymbol{M}(h)\boldsymbol{s} + \delta_h^2 \\ &\quad +2\boldsymbol{s}^T[\{I-\boldsymbol{W}(h)\mathrm{diag}(\nabla s_i)\}\boldsymbol{b}(h) - \{I-\boldsymbol{W}(\infty)\mathrm{diag}(\nabla s_i)\}\boldsymbol{b}(\infty)] \quad (20)\end{aligned}$$

where

$$\boldsymbol{M}(h) = 2\boldsymbol{W}(h) + \boldsymbol{W}(h)^T\mathrm{diag}(\nabla s_i)\boldsymbol{W}(h) - \boldsymbol{W}(\infty) \quad (21)$$

and

$$\delta_h^2 = \boldsymbol{b}(h)^T\mathrm{diag}_n(\nabla s_i)\boldsymbol{b}(h) - \boldsymbol{b}(\infty)^T\mathrm{diag}_n(\nabla s_i)\boldsymbol{b}(\infty).$$

Since $E(\boldsymbol{s}) = 0$, components in (20) has zero expectation so that

$$E\{\Lambda(h)\} \approx \mathrm{tr}\{\boldsymbol{M}(h)\mathrm{diag}_n(-\nabla s_i)\} + \delta^2.$$

Since under $H_0$ $\delta_h^2 \equiv 0$, the bias component steers the power of the test. Moreover $\mathrm{tr}\{\boldsymbol{M}(h)\mathrm{diag}_n(-\nabla s_i)\}$ gives the difference in the degree of model (2) fitted with bandwidth $h$ compared to model (1). Clearly, setting $h\to\infty$ this collapses to zero.

*One Step Estimate (9)*

Finally, we sketch some properties of the one-step estimate. By simple Taylor expansion and assuming bounded components in the model we get

$$\begin{aligned}\left\{\sum_{i=1}^{n} w_{li}\nabla s_i(\widehat{\beta}|\mathcal{R}_i)\right\}^{-1} &= \left\{\sum_{i=1}^{n} w_{li}\nabla s_i\right\}^{-1}\left\{1+O\left(\left\{\sum_{i=1}^{n} w_{li}\nabla s_i\right\}^{-1}\right)\right\} \\ &\approx \left\{\sum_{i=1}^{n} w_{li}\nabla s_i\right\}^{-1}\end{aligned}$$

where $O(\cdot)$ refers to vector valued orders. Again by Taylor expansion of (9) we get from (15)

$$\begin{aligned}
\widehat{\beta}_l^{(1)} &= \widehat{\beta}_0 - \left\{\sum_{i=1}^n w_{li}\nabla s_i\right\}^{-1}\left[\sum_{i=1}^n w_{li}\left\{s_i + \nabla s_i(\widehat{\beta}_0 - \beta_l + \beta_l - \beta_i) + O\{(\widehat{\beta}_l - \beta_i)^2\}\right\}\right] + \dots \\
&= \left\{\sum_{i=1}^n w_{li}\nabla s_i\right\}^{-1}\left\{\sum_{i=1}^n w_{li}s_i + b_{li}\right\} + \beta_l \qquad (22) \\
&\quad - \left\{\sum_{i=1}^n w_{li}\nabla s_i\right\}^{-1}\left[\sum_{i=1}^n w_{li}O\{(\widehat{\beta}_0 - \beta_i)^2\}\right] + \dots \\
&\approx \widehat{\beta}_l + O\{(\widehat{\beta}_l - \beta_l)^2\}. \qquad (23)
\end{aligned}$$

Approximation (23) follows from (22) by assuming $\beta(t)$ to be sufficiently smooth and applying standard smoothness arguments in the spirit of Cai & Sun (2002). From (23) it is seen that the one step estimate in first order approximation equals the fully iterated local partial likelihood estimate which motivates to use the one step estimate in the testing procedure.

## References


Abrahamowicz, M., MacKenzie, T., and Esdaile, J. M. (1996). Time-dependent hazard ratio: Modeling and hypothesis testing with application in lupus nephritis. *Journal of the American Statistical Association* **91**, 1432–1439.

Aerts, M., Claeskens, G., and Hart, J. D. (1999). Testing the fit of a parametric function. *Journal of the American Statistical Association.* **94**, 869–879.

Bowman, A. W. and Young, S. (1996). Graphical comparison of nonparametric curves. *Appl. Statist.* **45**, 83–98.

Breslow, N. (1974). Covariance analysis of censored survival data. *Biometrics* **30**, 89–99.

Breslow, N. E. (1972). Comment on "regression and life tables" by D. R. Cox. *Journal of the Royal Statistical Society, Series B* **34**, 216–217.

Cai, Z. and Sun, Y. (2002). Local linear estimation for time-dependent coefficients in cox's regression models. *Scandinavian Journal of Statistics*, (to appear).

le Cessie, S. and van Houwelingen, J. (1991). A goodness-of-fit test for binary regression models, based on smoothing methods. *Biometrics* **47**, 1267–1282.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, Series B* **B 34**, 187–220.

Davison, A. and Hinkley, A. (1997). *Bootstrap Methods and their Application.* Cambridge, UK: Cambridge University Press.

Fan, J., Gijbels, I., and King, M. (1997). Local likelihood and local partial likelihood in hazard regression. *Annals of Statist.* **25**, 1661–1690.

Grambsch, P. M. and Therneau, T. M. (1994). Proportional hazards tests and diagnostics based on weighted residuals (corr: 95v82 p668). *Biometrika* **81**, 515–526.

Gray, R. J. (1994). Spline-based tests in survival analysis. *Biometrics* **50**, 640–652.

Harbeck, N., Alt, U., Berger, U., Krüger, A., Thomssen, C., Jänicke, F., Höfler, H., Kates, R., and Schmitt, M. (2001). Prognostic impact of proteolytic factors (urokinase-type plasminogen activator, plasminogen activator inhibitor 1, and cathepsins b, d, and l) in primary breast cancer reflects effects of adjuvant systemic therapy. *Clinical Cancer Research* **7**, 2757–2764.

Harbeck, N., Thomssen, C., Berger, U., Ulm, K., K., R., Höfler, H., Jänicke, F., Graeff, H., and Schmitt, M. (1999). Invasion marker pai-1 remains strong prognostic factor after long-term follow-up both for primary cancer and following first relapse. *Breast Cancer Research and Treatment* **54**, 147–157.

Härdle, W. and Kneip, A. (1999). Testing a regression model when we have smooth alternatives in mind. *Scand. J. of Stat.* **26**, 221–238.

Härdle, W. and Mammen, E. (1993). Comparing nonparametric versus parametric regression fits. *Ann. Stat.* **21**, 1926–1947.

Hart, J. (1997). *Nonparametric Smoothing and Lack-of-Fit Tests.* Springer Verlag.

Hastie, T. and Tibshirani, R. (1990a). Exploring the nature of covariate effects in the proportional hazard model. *Biometrics* **46**, 1005–1016.

Hastie, T. and Tibshirani, R. (1990b). *Generalized Additive Models.* London: Chapman and Hall.

Hastie, T. and Tibshirani, R. (1993). Varying–coefficient models. *Journal of the Royal Statistical Society, Series B* **55**, 757–796.

Hertz-Picciotto, I. and Rockhill, B. (1997). Validity and efficiency of approximation methods for tied survival times in Cox regression. *Biometrics* **53**, 1151–1156.

Hess, K. R. (1994). Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. *Statistics in Medicine* **13**, 1045–1062.

Kauermann, G. and Tutz, G. (2001). Testing generalized linear and semiparametric models against smooth alternatives. *Journal of the Royal Statistical Society, Series B* **63**, 147 – 166.

Lin, D. Y. and Wei, L. J. (1991). Goodness-of-fit tests for the general Cox regression model. *Statistica Sinica* **1**, 1–17.

Moreau, T., O'Quigley, J., and Mesbah, M. (1985). A global goodness-of-fit statistic for the proportional hazards model. *Applied Statistics* **34**, 212–218.

Müller, H.-G. (1992). Goodness-of-fit diagnostics for regression models. *Scand. J. Statist.* **19**, 157–172.

O'Quigley, J. and Pessione, F. (1989). Score tests for homogeneity of regression effect in the proportional hazards model. *Biometrics* **45**, 135–144.

O'Sullivan, F. (1988). Nonparametric estimation of relative risk using splines and cross-validation. *SIAM J. Sci. Statist. Comput.* **9**, 531–542.

Sasieni, P. (1999). Cox regression model. In P. Armitage & T. Colton (Eds.), *Encyclopedia of Biostatistics*, Volume 1, pp. 1006–1020. New York: Wiley.

Verweij, P. and van Houwelingen, H. (1995). Time-dependent effects of fixed covariates in Cox regression. *Biometrics* **51**, 1550–1556.

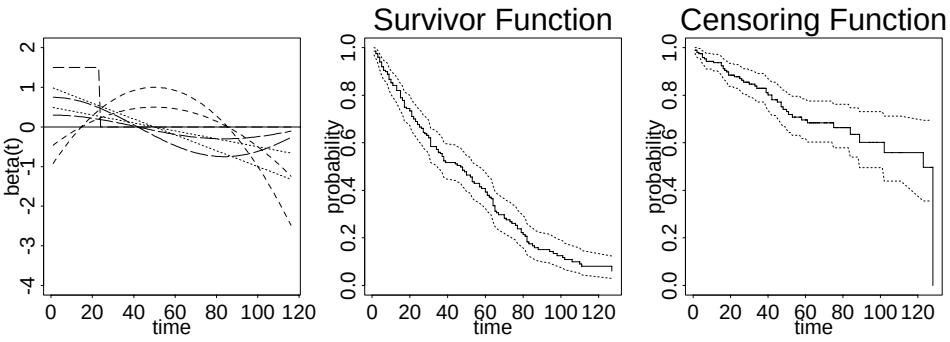


Figure 1: Different dynamic effect structures for $\beta(t)$ (left plot) and typical survival data resulting from the $H_0$ simulation setting represented by the estimated survivor function (middle plot) and the estimated censoring survivor function (right plot).

| | $\beta(t)$ | $\alpha = 0.05$ | $\alpha = 0.1$ |
|---|---|---|---|
| $H_0$ model: | $\beta(t) = 0$ | 5.2 | 8.5 |
| linear (flat): | $\beta(t) = 0.5 - t/100$ | 52.0 | 60.8 |
| quadratic (flat) | $\beta(t) = -0.5 + 4t/100 - 4t^2/10000$ | 46.4 | 57.2 |
| cosinus (weak) | $\beta(t) = 0.75 * cos(3t/40)$ | 24.6 | 39.8 |
| step function | $\beta(t) = 1.5I(t \leq 24)$ | 90.8 | 94.4 |
| linear (steep): | $\beta(t) = 1 - t/50$ | 89.2 | 94.8 |
| quadratic (steep) | $\beta(t) = -0.5 + 8t/100 - 8t^2/10000$ | 78.4 | 87.6 |
| cosinus (strong) | $\beta(t) = 1.5 * cos(3t/40)$ | 46.0 | 66.4 |

Table 1: Simulated rejection probability for testing the $H_0$ model (1) for different alternatives

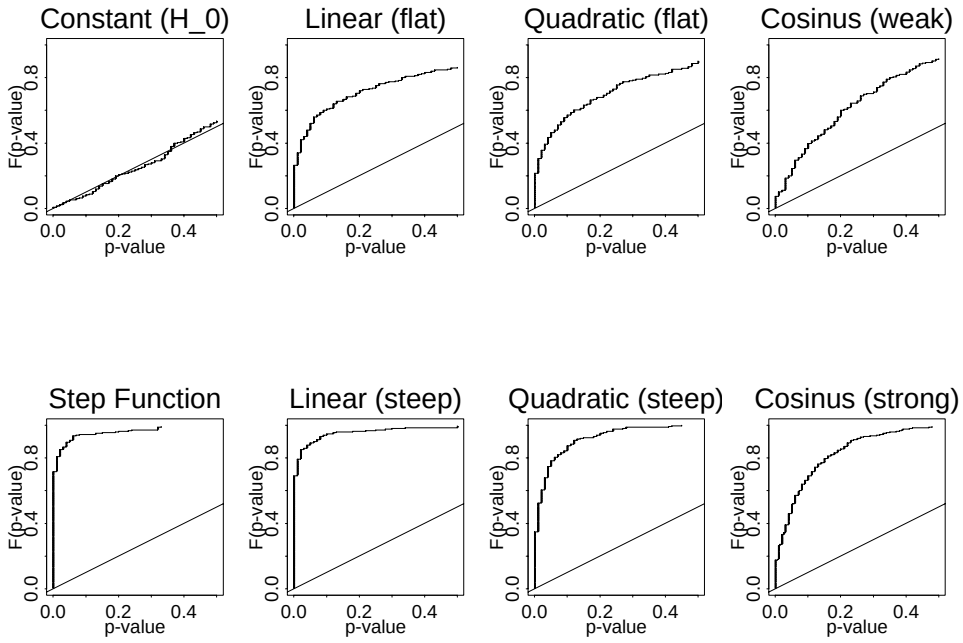


Figure 2: Distribution function of the $p - value$ for $\Lambda(h)$ for different simulation settings

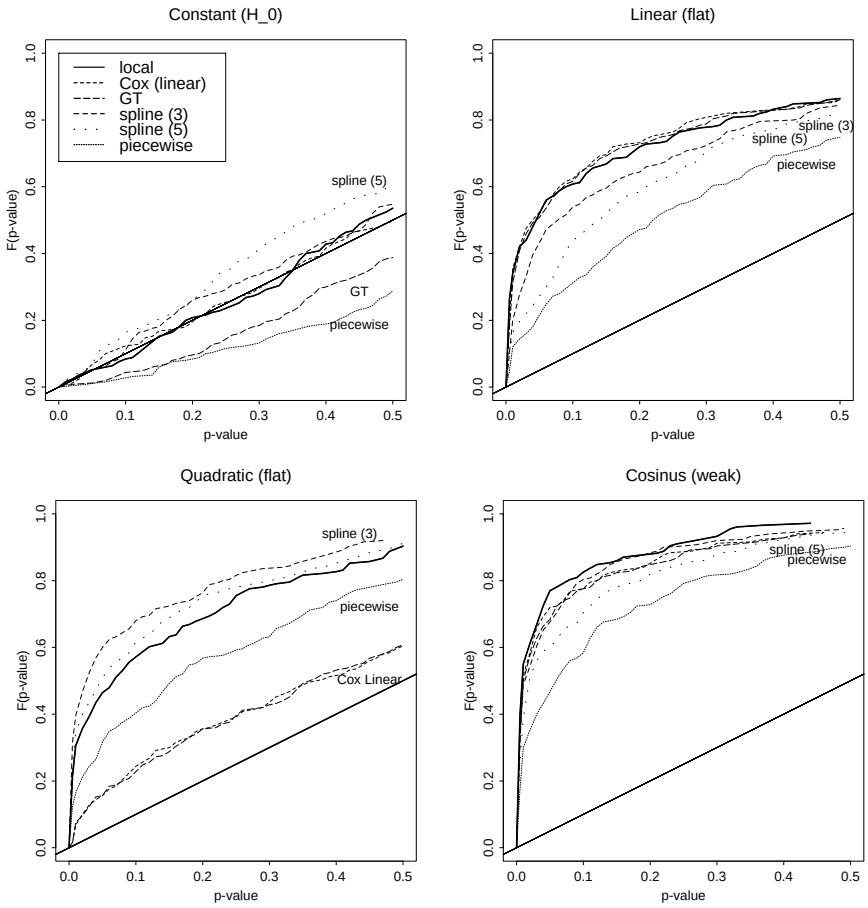


Figure 3: Empirical distribution of p-values for different tests in four simulations settings (thick solid line shows the local partial likelihood test).

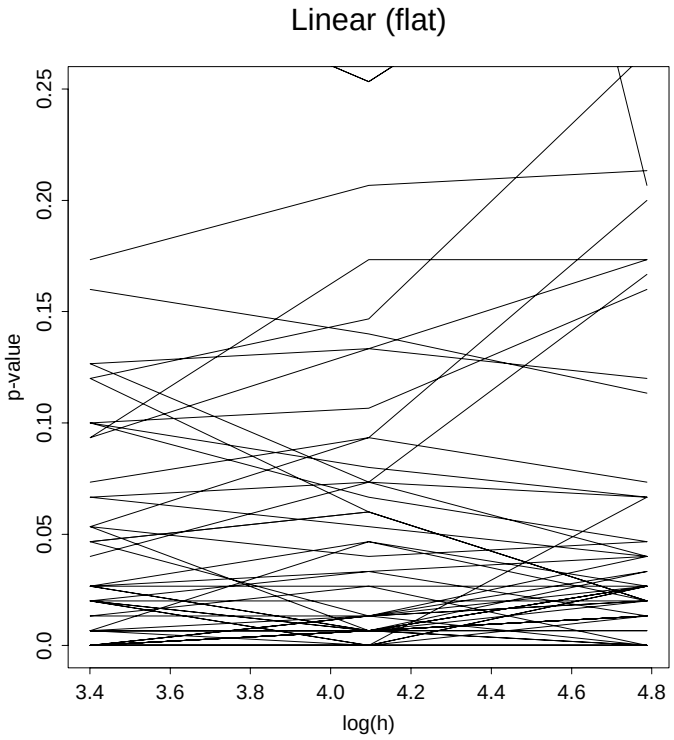


Figure 4: Trace of $p$-values for different bandwidths.

| simulation setting | decision | row bandwidth, column bandwidth $h = 30, h = 60$ accept | reject | $h = 30, h = 120$ accept | reject | $h = 60, h = 120$ accept | reject |
|---|---|---|---|---|---|---|---|
| $H_0$ | accept | 93 | 2 | 93 | 2 | 92 | 1 |
| | reject | 0 | 5 | 1 | 4 | 2 | 5 |
| | correlation | 0.95 | | 0.72 | | 0.84 | |
| linear | accept | 42 | 3 | 39 | 6 | 39 | 6 |
| | reject | 3 | 52 | 1 | 54 | 1 | 54 |
| | correlation | 0.98 | | 0.87 | | 0.92 | |
| quadratic | accept | 67 | 15 | 46 | 36 | 45 | 23 |
| | reject | 0 | 18 | 0 | 18 | 1 | 31 |
| | correlation | 0.88 | | 0.72 | | 0.92 | |

Table 2: Matching of same inferential conclusions (in percent based on 100 simulations) for tests based on 5 % significance level and correlation of corresponding $p$-values for different bandwidths.

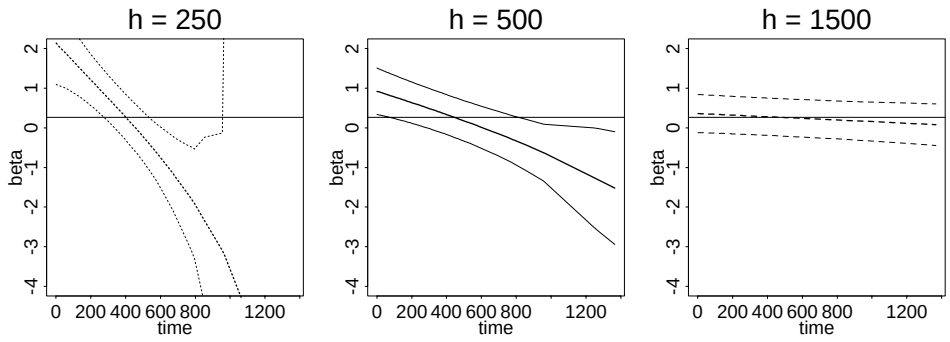


Figure 5: Local partial likelihood estimates and 95 % confidence intervals of the therapy effect on gastric cancer for three different bandwidths. (The parametric fit based on the PH-assumption is given as a reference line)

| | | | $p$-value for $H_0 : \beta(t) = \beta$ | | | |
|---|---|---|---|---|---|---|
| | | | Bandwidth | | | |
| Covariate | $\widehat{\beta}$ | sd($\widehat{\beta}$) | $h = 25$ | $h = 40$ | $h = 60$ | $h = 100$ |
| upa | 0.35 | 0.19 | 0.12 | 0.31 | 0.52 | 0.69 |
| pai | 0.74 | 0.21 | 0.02 | 0.01 | 0.01 | 0.02 |
| lypo | 1.12 | 0.21 | 0.36 | 0.8 | 0.86 | 0.83 |
| hormo | 0.16 | 0.22 | <0.01 | <0.01 | <0.01 | <0.01 |
| gradi | 0.25 | 0.20 | 0.19 | 0.17 | 0.13 | 0.09 |
| global | | | <0.01 | <0.01 | <0.01 | <0.01 |

Table 3: Parameter estimates and component-wise ( partial) and global p-values for the breast cancer data.

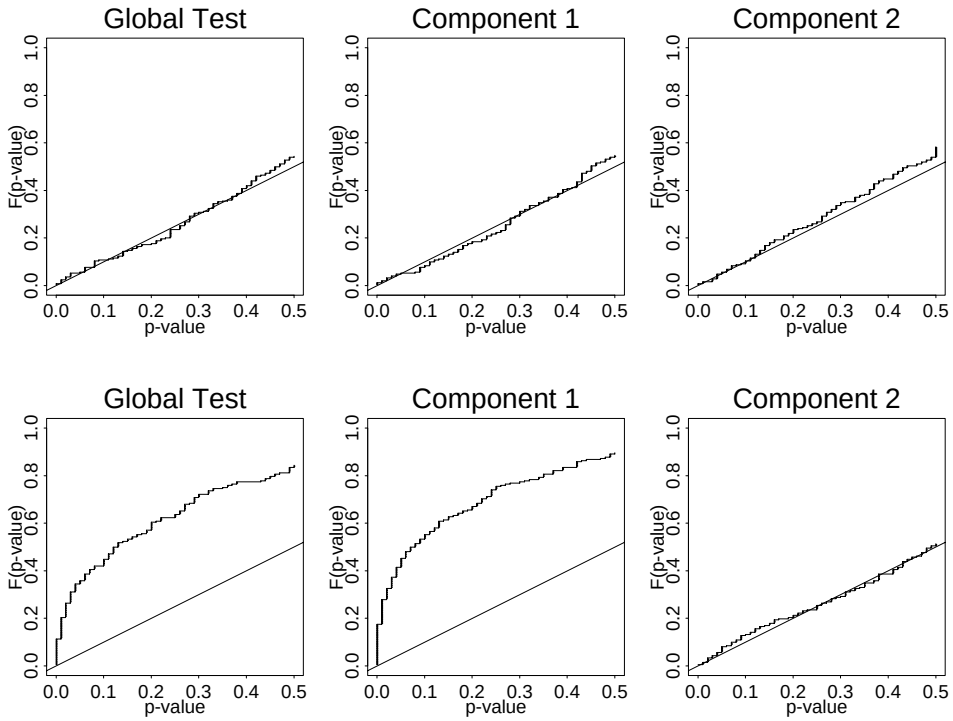


Figure 6: Probability function of p-value for $\Lambda(h)$, $\Lambda_1(h)$ and $\Lambda_2(h)$ for setting (a), upper row, and setting (b), bottom row.

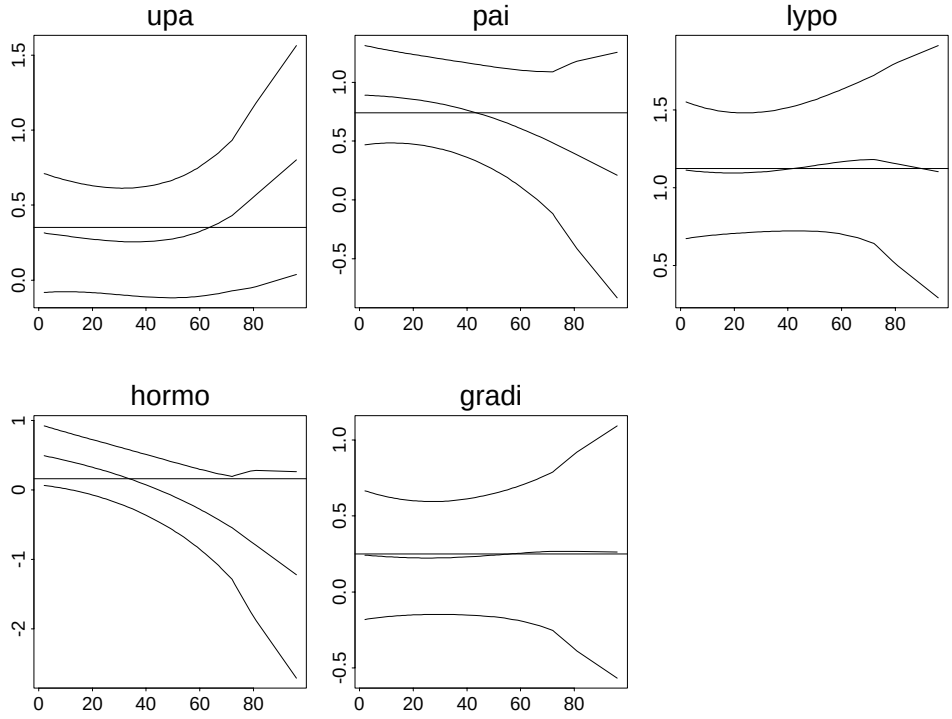


Figure 7: Varying coefficients for breast cancer data

Harbeck, Thomssen, Berger, Ulm, R., Jänicke, Graeff & Schmitt, 1999 or Harbeck, Alt, Berger, Krüger, Thomssen, Jänicke, Höfler, Kates & Schmitt, 2001)