University of California, Berkeley
U.C. Berkeley Division of Biostatistics Working Paper Series

Year 2002 *Paper* 117

Accelerated Hazards Model: Method, Theory and Applications

Ying Qing Chen[*] Nicholas P. Jewell[†]
Jingrong Yang[‡]

[*]Division of Biostatistics, School of Public Health, University of California, Berkeley, yqchen@stat.berkeley.edu

[†]Division of Biostatistics, School of Public Health, University of California, Berkeley

[‡]Division of Biostatistics, School of Public Health, University of California, Berkeley, jingrong@berkeley.edu

Accelerated Hazards Model: Method, Theory and Applications

Ying Qing Chen, Nicholas P. Jewell, and Jingrong Yang

Abstract

In an accelerated hazards model, the hazard functions of a failure time are related through the time scale-change, which is often a function of covariates and associated parameters. When the hazard functions have special properties, such as monotonicity in time, the parameters may be clinically meaningful in measuring a treatment effect. This paper reviews methodological and theoretical development of this model. Applications of the accelerated hazards model including sample size calculation in clinical trials, are also explored.

1 Introduction

Time-to-event or survival time data have been thoroughly studied over the past decades. The Cox proportional hazards model (Cox, 1972) is the most extensively used regression model in the analysis of such data. This model often assumes that a covariate effect is captured throught a proportionality constant between hazard functions, leaving the underlying hazard functions unspecified. Specifically, the Cox proportional hazards model for failure time T and associated covariates, Z, is

$$\lambda(t|Z) = \lambda_0(t)\exp(\beta^{\mathrm{T}}Z), \tag{1}$$

where $\lambda(\cdot)$ is the hazard function and β is a parameter vector of the same dimension as the covariate vector Z. Here, the superscript T denotes vector transpose. When censoring is present, the partial likelihood provides a simple and efficient way to estimate the parameter β. Within the framework of counting processes, asymptotic properties of estimators can be elegantly justified using martingale theory (Andersen and Gill, 1982). For references, see Kalbfleisch and Prentice (1980), Fleming and Harrington (1991) and Andersen, Borgan, Gill and Keiding (1993).

As noted in Lin and Ying (1997), the Cox model is a special case of more general hazard-based regression model:

$$\lambda(t|Z) = L\{\lambda_0(t), \beta^{\mathrm{T}}Z\}, \tag{2}$$

where $L(\cdot)$ is a known function. Other well-known examples of (2) include the additive hazards model (Lin and Ying, 1994)

$$\lambda(t|Z) = \lambda_0(t) + \beta^{\mathrm{T}}Z, \tag{3}$$

and the accelerated failure time model (Kalbfleisch and Prentice, 1980)

$$\lambda(t|Z) = \lambda_0\{t\exp(\beta^{\mathrm{T}}Z)\}\exp(\beta^{\mathrm{T}}Z). \tag{4}$$

Further discussion of the merits of (2) can be found in Lin and Ying (1997).

The accelerated hazards model (Chen and Wang, 2000a) is also a special case of (2):

$$\lambda(t|Z) = \lambda_0\{t\exp(\beta^{\mathrm{T}}Z)\}. \tag{5}$$

In this model, $\exp(\beta^{\mathrm{T}} Z)$ characterizes how the covariate Z alters the time scale of the underlying hazard progression, and is termed the hazard progression time ratio. Whether $\beta > 0$ or $\beta < 0$ reflects the direction of the alteration: acceleration or deceleration, respectively. For instance, assume that the covariate Z takes the value of 0 (control) or 1 (treatment). If β is $-\log 2$, the model claims that the hazard of the treatment group progresses in half the time as those in the control group. If β is $\log 2$, the hazard of the treatment group progresses in twice the time as those in the control group. If $\beta = 0$, it means there is no difference between the two groups in hazard progression.

In the rest of this article, we review and summarize methodological and theoretical developments of the accelerated hazards model in estimation and efficiency evaluation. Some techniques are presented to check for model adequacy. Several extensions of the model are introduced. Practical implementation of the model is also discussed.

2 Estimation

Let non-negative random variables T and C stand for the failure and censoring times, respectively, and let Z denote the p-vector covariate. Conditional on Z, T and C are assumed to be independent. We observe $X_i = \min(T_i, C_i)$, $\Delta_i = I(T_i \le C_i)$ and Z_i, for $i = 0, 1, \ldots, n$. Here $I(\cdot)$ is the indicator function taking the value of 1 if the condition in $(\cdot)$ is satisfied and 0 otherwise. For the ith individual, denote $N_i(t) = I(X_i \le t, \Delta_i = 1)$ and $Y_i(t) = I(X_i \ge t)$. Define the filtration

$$\mathcal{F}_t = \sigma\{N_i\{t\exp(-\beta_0^{\mathrm{T}} Z_i)\}, Y_i\{t\exp(-\beta_0^{\mathrm{T}} Z_i)\}, Z_i; i = 1, 2, \ldots, n\}.$$

Then, assuming the accelerated hazards model (5), for the ith individual,

$$\begin{aligned} E[dN_i\{t\exp(-\beta_0^{\mathrm{T}} Z_i)\}|\mathcal{F}_{t-}] &= Y_i\{t\exp(-\beta_0^{\mathrm{T}} Z_i)\} d\Lambda_i\{t\exp(-\beta_0^{\mathrm{T}} Z_i)\} \\ &= Y_i\{t\exp(-\beta_0^{\mathrm{T}} Z_i)\} \exp(-\beta_0^{\mathrm{T}} Z_i) d\Lambda_0(t), \end{aligned}$$

where $\Lambda(\cdot) = \int_0^{\cdot} \lambda(u)du$ is the cumulative hazard function and β_0 is the true value of parameter β. Let

$$M_i(t; \beta, \lambda_0) = N_i\{t\exp(-\beta^{\mathrm{T}} Z_i)\} - \int_0^t Y_i\{t\exp(-\beta^{\mathrm{T}} Z_i)\} \exp(-\beta^{\mathrm{T}} Z_i) d\Lambda_0(t).$$

Then $E\{M_i(t)\} = 0$ and $M_i(t; \beta_0, \Lambda_0)$ are martingales with respect to $\mathcal{F}$. It is therefore reasonable to estimate β through

$$\sum_{i=1}^{n} \int_0^\infty G(t, Z; \beta) dM_i(t; \beta, \Lambda_0) = 0, \tag{6}$$

where $G(t, Z; \beta)$ is a known two-dimensional weight function, and $G(t, Z; \beta_0)$ is measurable with respect to the filtration $\mathcal{F}_t$. For example, one possible choice of G is to let $G(t, Z; \beta) = (1, Z)^{\mathrm{T}}$. Then the above equations become

$$\sum_{i=1}^{n} \int_0^\infty dM_i(t; \beta, \Lambda_0) = 0, \text{ and } \sum_{i=1}^{n} \int_0^\infty Z_i dM_i(t; \beta, \Lambda_0) = 0. \tag{7}$$

Furthermore, by replacing $\Lambda_0(t)$ with its Breslow-type of estimator

$$\hat{\Lambda}_0(t, \beta) = \int_0^t \frac{\sum_{i=1}^n dN_i\{t \exp(-\beta^{\mathrm{T}} Z_i)\}}{\sum_{i=1}^n Y_i\{t \exp(-\beta^{\mathrm{T}} Z_i)\} \exp(-\beta^{\mathrm{T}} Z_i)},$$

and doing some algebraic manipulation, we can use the following unbiased estimating functions for parameter estimation:

$$S(\beta) = \sum_{i=1}^{n} \int_0^\infty \{Z_i - \bar{Z}(t, \beta)\} dN_i\{t \exp(-\beta^{\mathrm{T}} Z_i)\}, \tag{8}$$

where

$$\bar{Z}(t, \beta) = \frac{\sum_{j=1}^n Y_j\{t \exp(-\beta^{\mathrm{T}} Z_j)\} \exp(-\beta^{\mathrm{T}} Z_j) Z_j}{\sum_{j=1}^n Y_j\{t \exp(-\beta^{\mathrm{T}} Z_j)\} \exp(-\beta^{\mathrm{T}} Z_j)}. \tag{9}$$

A weighted version of (8) is

$$S^W(\beta) = \sum_{i=1}^{n} \int_0^\infty W(t, \beta)\{Z_i - \bar{Z}(t, \beta)\} dN_i\{t \exp(-\beta Z_i)\}, \tag{10}$$

then the weight function $W(t, \beta_0)$ is left-continuous, non-negative and measurable with respect to $\mathcal{F}_t$, and $n^{-1} W(t, \beta_0)$ is assumed to uniformly converge to a nonrandom function $w(t, \beta_0)$. When $\hat{\beta}$ is available, the baseline cumulative hazard function can be estimated by $\hat{\Lambda}_0(t, \hat{\beta})$:

$$\int_0^t \frac{d \sum_{i=1}^n N_i\{u \exp(-\hat{\beta}^{\mathrm{T}} Z_i)\}}{\sum_{i=1}^n Y_i\{u \exp(-\hat{\beta}^{\mathrm{T}} Z_i)\} \exp(-\hat{\beta}^{\mathrm{T}} Z_i)}. \tag{11}$$

3 Asymptotic results

Since $S(\beta)$ is not a continuous function of β, a unique solution to $S(\beta) = 0$ is very unlikely. Following the definitions used in Tsiatis (1990) or Wei, et al. (1990) for rank estimation of the accelerated failure time model, the estimator of $\hat{\beta}$ can be defined as (i) zero-crossing of $S(\beta)$ such that $S(\hat{\beta}-)S(\hat{\beta}+) \leq 0$ when β is one-dimensional, or (ii) the minima of $\|S(\beta)\|$.

Let $S(t;\beta)$ be

$$S(t;\beta) = \sum_{i=1}^{n} \int_0^t \{Z_i - \bar{Z}(t,\beta)\} dN_i\{t \exp(-\beta^{\mathrm{T}} Z_i)\}.$$

Then $n^{-1/2} S(t, \beta_0)$ converges weakly to a zero-mean Gaussian process with variance function, $\Sigma(t, \beta_0)$, say, which can be consistently estimated by

$$n^{-1} \sum_{i=1}^{n} \int_0^t \{Z_i - \bar{Z}(t,\beta)\}^{\otimes 2} dN_i\{t \exp(-\beta^{\mathrm{T}} Z_i)\},$$

where, for a vector v, $v^{\otimes 2}$ denotes vv^{T}.

Suppose the following regularity conditions hold:

1. the covariates, Z_i, are uniformly bounded;
2. the censoring times, C_i, have uniformly bounded densities;
3. λ_0 has a bounded second derivative.

Following the similar arguments in Ying (1993), We can show that $\hat{\beta}$ is consistent and

$$n^{1/2}(\hat{\beta} - \beta_0) \xrightarrow{d} N(0, D^{-1}\Sigma(D^{-1})^{\mathrm{T}}),$$

where

$$D = \int_0^{\infty} E\left[Y_1\{t \exp(-\beta^{\mathrm{T}} Z_1)\} \exp(-\beta^{\mathrm{T}} Z_1)(Z_1 - \bar{Z})\right] d\{\lambda_0(t)t\}.$$

Using the techniques and the asymptotic linearity property as described in Appendix 3 of Chen and Jewell (2001), it is also true that $n^{1/2}\{\hat{\Lambda}_0(t,\hat{\beta}) - \Lambda_0(t,\beta_0)\}$ converges weakly to a zero-mean Gaussian process.

4 Efficiency considerations

The estimating equations proposed in the preceding section are *ad hoc*. The approaches in Lai and Ying (1992), Lin and Ying (1994) and Chen and Jewell (2001) can be followed to study a special parametric sub-family for the semiparametric information bound. Let such parametric sub-family be

$$\lambda(t|Z,\alpha,\beta) = \alpha h\{t\exp(\beta Z)\} + \lambda_0\{t\exp(\beta Z)\}, \tag{12}$$

where α and β are unknown parameters, α_0 and β_0 are their true values, respectively, and $\lambda_0(\cdot)$ and h are fixed functions. Consider the log likelihood function for (α,β) in (12):

$$l(\alpha,\beta) = \sum_{i=1}^{n}\left\{\int_0^\infty \log\lambda(t|Z_i,\alpha,\beta)dN(t|Z_i) - \int_0^\infty Y(t|Z_i)\lambda(t|Z_i,\alpha,\beta)dt\right\}$$

$$= \sum_{i=1}^{n}\left(\int_0^\infty \log[\alpha h\{t\exp(\beta Z_i)\} + \lambda_0\{t\exp(\beta_1 Z_i)\}]dN_i(t) - \int_0^\infty Y_i(t)[\alpha h\{t\exp(\beta Z_i)\} + \lambda_0\{t\exp(\beta Z_i)\}]dt\right).$$

Denote the Fisher Information matrix of $I_h(\alpha,\beta)$ at $\alpha = 0$ and $\beta = \beta_0$ by

$$I_h(0,\beta_0) = \begin{pmatrix} I_{\alpha\alpha}(h) & I_{\alpha\beta}(h) \\ I_{\beta\alpha}(h) & I_{\beta\beta}(h) \end{pmatrix}. \tag{13}$$

Then the semiparametric information bound of β can be calculated by exhausting all the choices of h in (12) at $\alpha = \alpha_0$ and $\beta = \beta_0$, yielding

$$I_{\beta\beta}(h) - I_{\beta\alpha}(h)I_{\alpha\alpha}^{-1}(h)I_{\alpha\beta}(h), \tag{14}$$

by the Cramér-Rao inequality. Straightforward calculation shows that the set of semiparametrically efficient estimating equations for the parameter β_0 is given by

$$S_{\text{opt}}(\beta) = \sum_{i=1}^{n}\int_0^\infty \{\lambda_0^{(1)}(t)t/\lambda_0(t)\}\{Z_i - \bar{Z}(t;\beta)\}dN_i\{t\exp(-\beta_1 Z_i)\} = 0, \tag{15}$$

where $\lambda_0^{(1)}(t)$ is the first derivative of $\lambda_0(t)$ with respect to t. It is possible to attain the semiparametric efficiency bound, although the use of S_{opt} may be limited in practice because the optimal weight involves the unknown baseline hazard function.

5 Model adequacy

The accelerated hazards models, together with other classes of models, providing flexibilityin modeling survival data with censored observations. However, careful assessment of adequacy of this model is a critical issue. As demonstrated in Chen (2001), two test statistics are available for checking model adequacy.

Kolmogorov-Smirnov Test

A Kolmogorov-Smirnov test can be used when Z is a binary covariate in the accelerated hazards regression model. As studied in Chen (2001), the following Kolmogorov-Smirnov type of test statistic can be used:

$$T_{\text{KS}}(\hat{\beta}) = \|D(t,\hat{\beta})\| \stackrel{\text{def}}{=} \sup_{0 \le t < \infty} |D(t,\hat{\beta})|, \tag{16}$$

where

$$D(t,\hat{\beta}) = n^{-1/2} \int_0^t Q(u,\hat{\beta}) \left[\exp(\hat{\beta}) d\hat{\Lambda}_1\{\exp(-\hat{\beta})\} - \hat{\Lambda}_0(u)\right]. \tag{17}$$

Here, $Q(t,\beta_0)$ is measurable with respect to $\mathcal{F}_t$, and $n^{-1}Q(t,\beta_0)$ has the limiting function of $q(t,\beta_0)$, as $n \to \infty$. Choices for Q include:

$$Q_{\text{LR}}(u,\beta) = \frac{Y_0(u)Y_1\{u\exp(-\beta)\}\exp(-\beta)}{Y_0(u) + Y_1\{u\exp(-\beta)\}\exp(-\beta)}$$

of the Log-rank test;

$$Q_{\text{GE}}(u,\beta) = n^{-1}Y_0(u)Y_1\{u\exp(-\beta)\}\exp(-\beta)$$

of the Gehan's test;

$$Q_{\text{PP}}(u,\beta) = \hat{S}(u-,\beta)\frac{Y_0(u)Y_1\{u\exp(-\beta)\}\exp(-\beta)}{Y_0(u) + Y_1\{u\exp(-\beta)\}\exp(-\beta)}$$

of the Peto-Prentice generalized Wilcoxon statistics, where $\hat{S}(u-,\beta)$ is the left-continuous version of the Kaplan-Meier estimate based on the pooled sample of $\{(X_i\exp(\beta Z_i),\Delta_i), i = 1,2,\ldots,n\}$; and

$$Q_{\text{HF}}(u,\beta) = \{\hat{S}(u-,\beta)\}^{\rho}\frac{Y_0(u)Y_1\{u\exp(-\beta)\}\exp(-\beta)}{Y_0(u) + Y_1\{u\exp(-\beta)\}\exp(-\beta)}$$

of the general class of tests by Harrington and Fleming (1982).

An unexpectedly large value of $T_{\text{KS}}(\hat{\beta})$ leads to rejection of H_0, the null hypothesis that the accelerated hazards model is adequate. T_{KS} is consistent against any general alternative hypothesis when the accelerated hazards model is inadequate and hence omnibus. The selection of critical values of T_{KS} is discussed in the Appendix of Chen (2001).

Gill-Schumacher Test

For a general covariate vector Z, let $\hat{\beta}_1$ and $\hat{\beta}_2$ be the solutions of $S_1(\beta) = 0$ and $S_2(\beta) = 0$ with different choices of weight functions of W_1 and W_2, respectively. When the accelerated hazards model is true, the difference between $\hat{\beta}_2$ and $\hat{\beta}_1$ should be small and therefore a Wald-type statistic, based on $\hat{\beta}_2 - \hat{\beta}_1$, can be used to test the model adequacy. Following similar arguments in Wei, et al. (1990) for assessing adequacy of the accelerated failure time model, we see that $n^{-1/2}(S_1(\beta_0), S_2(\beta_0))^{\text{T}}$ is asymptotically joint Normal with covariance matrix, the limit of

$$n^{-1}\begin{bmatrix} V_{11}(\beta_0) & V_{12}(\beta_0) \\ V_{21}(\beta_0) & V_{22}(\beta_0) \end{bmatrix},$$

where

$$V_{11}(\beta_0) = \sum_{i=1}^{n}\int_0^\infty W_1^2(Z_i - \bar{Z})^{\otimes 2} dN_i\{t\exp(-\beta_0^{\text{T}} Z_i)\}$$

$$V_{12}(\beta_0) = V_{21}(\beta_0) = \sum_{i=1}^{n}\int_0^\infty W_1 W_2(Z_i - \bar{Z})^{\otimes 2} dN_i\{t\exp(-\beta_0^{\text{T}} Z_i)\}$$

$$V_{21}(\beta_0) = \sum_{i=1}^{n}\int_0^\infty W_2^2(Z_i - \bar{Z})^{\otimes 2} dN_i\{t\exp(-\beta_0^{\text{T}} Z_i)\}.$$

Therefore the following statistic,

$$T_{\text{GS}} = \min_{\beta\in U(\hat{\beta}_1)}\left\{\begin{bmatrix} S_1(\beta) \\ S_2(\beta + \hat{\beta}_2 - \hat{\beta}_1) \end{bmatrix}'\begin{bmatrix} V_{11}(\hat{\beta}_1) & V_{12}(\hat{\beta}_1) \\ V_{21}(\hat{\beta}_1) & V_{22}(\hat{\beta}_1) \end{bmatrix}^{-1}\begin{bmatrix} S_1(\beta) \\ S_2(\beta + \hat{\beta}_2 - \hat{\beta}_1) \end{bmatrix}\right\},$$

is asymptotically equivalent to a Wald-type statistic for testing $\hat{\beta}_2 - \hat{\beta}_1$, and follows a χ_p^2-distribution asymptotically, where p is dimension of the β-vector.

6 Extensions

The accelerated hazards model can be extended in various directions that we now briefly discuss.

Scale change function

The treatment need not to be limited to changing the time scale linearly. For example, the following scale change function can be used:

$$\lambda(t|Z) = \lambda_0[\exp\{\eta(t, Z; \beta)\}], \tag{18}$$

where the known positive parametric function $\eta(\cdot)$ is monotonically increasing in t. For example, with binary Z and $\eta(t, Z; \beta) = (1 - Z)t + Z\beta_0 t^2$, the treatment changes the time scale quadratically. If $\eta(t, z; \beta) = t(\beta_0^Z + \beta_1^Z t)/(1 + \beta_1^Z t)$, then $\eta(t, Z = 1; \beta)/(\beta_0 t) \to 1$ as $t \to 0$ and $\eta(t, z = 1; \beta)/t \to 1$ as $t \to \infty$. Therefore β_0 characterizes the acceleration rate in the early time period while β_1 describes how fast the acceleration effect is dampened as time progresses.

General model I

In some situations, not all covariates are considered to alter the time scale of hazard function. Such covariates, for example, may include age, gender and social economic status, which may still have proportional influence on the baseline hazard function. Let $Z = (Z_1, Z_2, \ldots, Z_p)^{\mathrm{T}} = (Z_{p_1}, Z_{p_2})^{\mathrm{T}}$, where Z_{p_1} are the first p_1 covariates and Z_{p_2} the rest of the p covariates. Then a variation of the accelerated hazards model for this situation is

$$\lambda(t|Z) = \lambda_0\{t \exp(\beta_{p_1}^{\mathrm{T}} Z_{p_1})\} \exp(\beta_{p_2}^{\mathrm{T}} Z_{p_2}) \tag{19}$$

where $\beta = (\beta_1, \beta_2, \ldots, \beta_p) = (\beta_{p_1}, \beta_{p_2})$. This model assumes that some covariates influence the baseline hazard function proportionally, while the baseline hazard function itself is accelerated or decelerated according to another set of covariates. Therefore, β_{p_1} is interpreted as the covariate effect on acceleration/deceleration of the baseline hazard progression as if Z_{p_2} is held constant; similarly, β_{p_2} is interpreted as the covariate effect of Z_{p_2} on relative hazards while the other covariates Z_{p_1} are held constant.

General model II

When all the covariates are considered to have impact on both the time scale and the magnitude of the baseline hazard function, it is natural to use a more general hazards regression model such as

$$\lambda(t|Z) = \lambda_0[t\exp(\beta_1^{\mathrm{T}} Z)]\exp(\beta_2^{\mathrm{T}} Z). \tag{20}$$

It is not difficult to see that (20) includes the Cox proportional hazards model, the accelerated failure time model and the accelerated hazards model. Namely, when $\beta_1 = 0$, (20) becomes the Cox proportional hazards model with a proportionality constant of $\exp(\beta_2)$; when $\beta_1 = \beta_2$, (20) becomes the accelerated failure time model with a time scale change of $\exp(\beta_1)$ or $\exp(\beta_2)$ in the survival functions; when $\beta_2 = 0$, (20) becomes the accelerated hazards model with the hazards progression time ratio of $\exp(\beta_{10})$. Therefore, (20) provides an approach to judge which one of the three models may well fit the actual data. When the underlying distribution is Weibull, all these three classes of models, and the general model in (20) coincide, as shown in Proposition 1 in Chen and Jewell (2001),

7 Implementation and Application

Although the semiparametrically efficient estimating functions of $S_{\mathrm{opt}}(\beta)$ are available, using them in practical analysis may be difficult, simply because of the difficulty in estimating the ratio of $\lambda_0^{(1)}(t)$ and $\lambda_0(t)$ from the observed data. Lin and Ying (1994) proposed the so-called "sample splitting" technique for constructing efficient estimators for the additive-multiplicative hazards model, that may be extended to the current situation. In general, however, if users are willing to sacrifice some efficiency, there are many other simpler weight functions, G, such as the Gehan's weight function. Simulation studies on different weight functions can be found in Chen and Jewell (2001).

Solving the estimating equations is challenging, because the estimating functions are not smooth. Standard numerical approaches, such as the Newton-Raphson algorithm, are inefficient. When the covariates are of low dimension, direct grid search or the bisection method can be used; in high dimensions, random search techniques such as "simulated annealing" might be more efficient. In cases of moderate dimension, the "recursive bisection" method by Huang (2002) is recommended to search for solutions. The rationale of this recursive method is simple: for example, in the kth step, suppose we know how to solve for $(\beta_1, \beta_2, \ldots, \beta_{k-1})$, then equipped with a one-dimensional bisection algorithm, we should be able to solve for β_k. From our experience, the computing time of this recursive method is modest and acceptable.

It appears straightforward to estimate the variance of $\hat{\beta}$ by simply replacing D with any consistent estimators, $\hat{D}$ and Σ with $\hat{\Sigma}$, respectively. However, since D involves the unknown baseline hazard function $\lambda_0(t)$ and furthermore its derivative, it is extremely difficult to obtain $\hat{D}$ directly, although there are several approaches of which we can take advantage. For example, when the sample size is large enough, the nonparametric kernel density estimation suggested by Tsiatis (1990) can be used; alternatively, a computing-intensive resampling algorithm of Parzen, et al. (1994) can be implemented to approximate the variance-covariance matrix. An approach based on numerical differences by Huang (2002) may be used. Details can be found in Chen and Jewell (2001).

Examples of the accelerated hazards model applied to specific data sets can be found in Chen and Wang (2000a & 2000b). Simulation studies can be conducted to compute sample sizes for use in planning clinical trials to detect the alternative of the accelerated hazards models. For we hereby describe simulation studies that have been conducted on (5), with the baseline hazard function of standard log-logistic: $\lambda_0(t) = 1/(1+t)$). A two-sample semiparametric score test statistic based on $S(\beta)$ is used to generate power curves for different combinations of β, censoring percentages and sample sizes. Simulations were conducted using 1,000 iterations. Figures 1 and 2 show the power calculations using contour plots and the three-dimensional graphs, respectively.

[Figure 1. and 2. about here]

8 Final remarks

The accelerated hazards model carries some unique features of its own. First, it treats hazard function as process of hazard progression over time and therefore the parameter in the model has an interpretation of hazard progression acceleration/deceleration. Furthermore, when baseline hazard function $\lambda_0(t)$ is monotone, the parameter identifies certain meaningful treatment effect (Chen and Wang, 2000b). Second, in contrast to the Cox model and the additive hazards model, the accelerated hazards model is not necessarily restricted to constant proportionality or constant additivity. This property is similar to the accelerated failure time model, which may lead to more flexibility in empirical sense to certain types of data. Therefore, the accelerated hazards model serves as a valuable supplement to the available hazards regression models.

When the assumptions of other hazard regression models are not reasonable in practice, the accelerated hazard model offers an alternative or supplement. This model may be more suitable in randomized clinical trials when the treatment effect is assumed to affect a time scale change between hazard functions. It is not limited to the crossovers in the hazard functions or survival functions. This allows the use of a single parameter to reflect a complex phenomenon, which otherwise requires time-dependent covariate effect in other models.

The accelerated hazards model is subject to an identifiability condition. That is, it is not identifiable when the baseline hazard function is constant, i.e., when the underlying distribution is exponential. This identifiability condition needs to be checked before implementing the accelerated hazards model, otherwise the variance estimates of the parameter of interest will be unreasonably large and subsequent inference becomes less meaningful.

References

Andersen, P. K., Borgan, Ø., Gill, R. D. and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer, New York.

Andersen, P. K. and Gill R. D. (1982). Cox's Regression Model for Counting Processes: A Large Sample Study. *Ann. Statist.* 4, 1100-20.

Chen, Y. Q. (2001) Accelerated hazards regression model and its adequacy for censored survival data. *Biometrics* 57, 853-860.

Chen, Y. Q. and Jewell, N. P. (2001) On a general class of semiparametric hazards regression models. *Biometrika* 88, 687-702.

Chen, Y. Q. and Wang, M.-C. (2000a). Analysis of accelerated hazards model. *J. of Amer. Statist. Asso.* 95, 608-618.

Chen, Y. Q. and Wang, M.-C. (2000b). Estimating a treatment effect with the accelerated hazards models. *Controlled Clinical Trials* 21, 369-380.

Cox, D. R. (1972). Regression Models and Life-Tables (with discussion). *J. R. Statist. Soc.* Ser. B 33, 187-220.

Fleming, T. and Harrington, D. (1991). *Counting Processes and Survival Analysis.* John Wiley, New York.

Harrington, D. P. and Fleming, T. R. (1982). A class of rank test procedures for censored survival data. *Biometrika* 69, 133-143.

Huang, Y. J. (2002). Calibration regression of censored lifetime medical cost. *J. Amer. Statist. Asso.* 97, 318-327.

Kalbfleisch, J. D. and Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data.* John Wiley, New York.

Lai, T. Z. and Ying, Z. (1992). Linear rank statistics in regression analysis with censored or truncated Data. *J. Multi. Anal.* 40, 13-45.

Lin, D. Y. and Ying, Z. (1994). Semiparametric analysis of the additive risk model. *Biometrika* 81, 61-71.

Lin, D. Y. and Ying (1997). Additive hazards regression models for survival data. In: D. Y. Lin and T. R. Fleming, eds., *Proceedings of the First Seattle Symposium in Biostatistics: Survival Analysis.* Springer, New York.

Parzen, M.I., Wei, L.J. and Ying, Z. (1994). A resampling method based on pivotal estimating functions. *Biometrika* 81, 341-350.

Tsiatis, A. A. (1990). Estimating regression parameters using linear rank tests for censored data. *Ann. Statist.*, 18, 354-72

Wei, L. J., Ying, Z. and Lin, D. Y. (1990). Linear Regression Analysis of Censored Survival Data Based on Rank Tests, *Biometrika* 77, 845-51.

Ying, Z. (1993). A large sample study of rank estimation for censored regression data. *Ann. Statist.* 21, 76-99.

Figure 1: Contour plots of simulated power curves of the accelerated hazards model. Four censoring percentages are used in graphs: (A) 0% (B) 10% (C) 25% (D) 50%. The marked numbers in the dashed lines are corresponding powers.

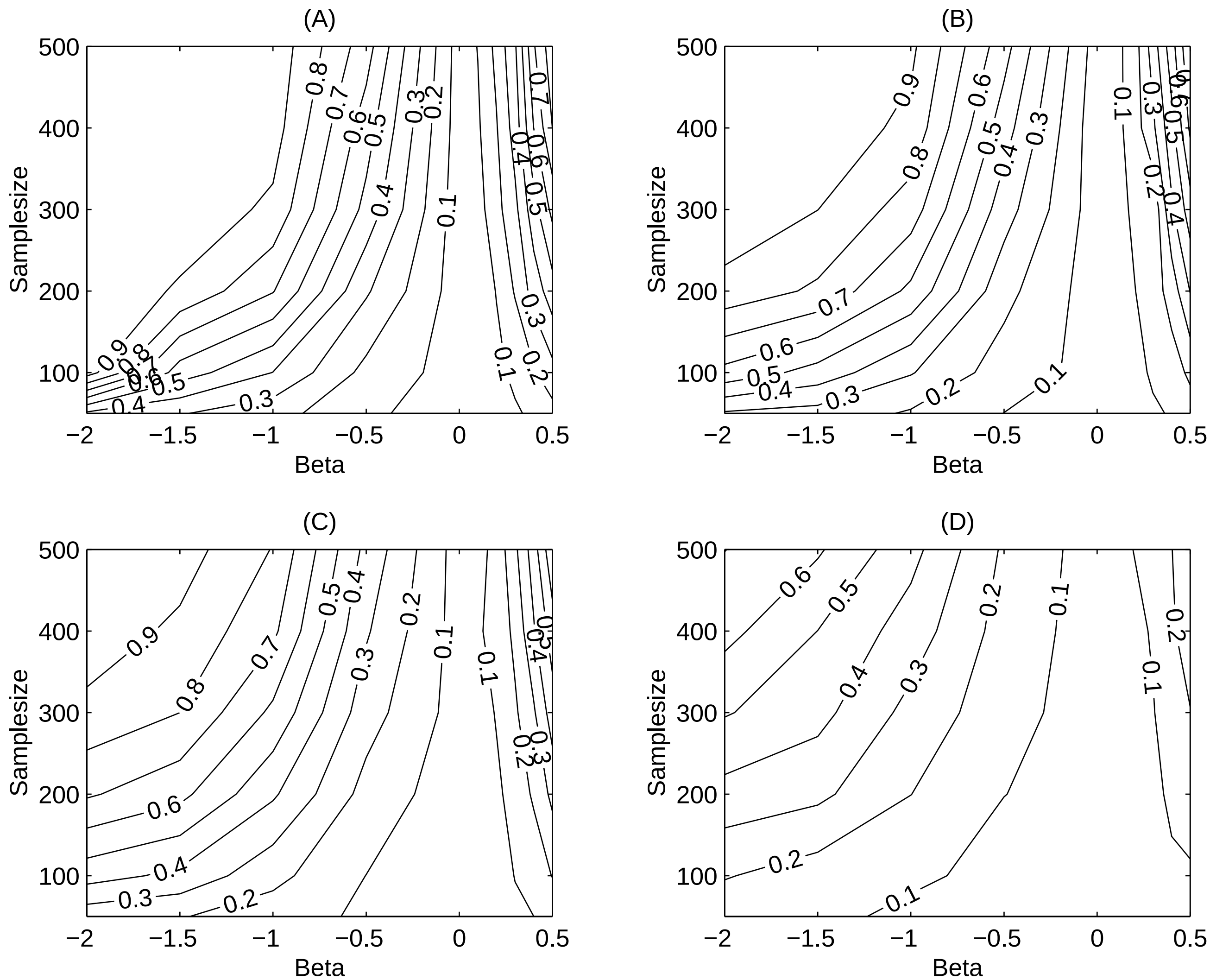

Figure 2: Three dimensional plots of simulated power curves of the accelerated hazards model. Four censoring percentages are used in graphs: (A) 0% (B) 10% (C) 25% (D) 50%.