Trend-constrained corrected score for proportional hazards model with covariate measurement error

Yijian Huang, Emory University
M Zhu, AbbVie

Journal Title: Contemporary Clinical Trials Communications
Volume: Volume 1, Number
Publisher: (publisher) | 2015-10-30, Pages 5-16
Type of Work: Article
Publisher DOI: 10.1016/j.conctc.2015.08.001
Permanent URL: https://pid.emory.edu/ark:/25593/s9q44

Final published version: http://dx.doi.org/10.1016/j.conctc.2015.08.001

Accessed June 1, 2018 12:43 AM EDT
Trend-constrained corrected score for proportional hazards model with covariate measurement error

Ming Zhu\textsuperscript{a},*, Yijian Huang\textsuperscript{b}

\textsuperscript{a} Abbvie Inc, 1 North Waukegan Rd, North Chicago, IL 60064, USA
\textsuperscript{b} Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA

ABSTRACT

In many medical research studies, survival time is typically the primary outcome of interest. The Cox proportional hazards model is the most popular method to investigate the relationship between covariates and possibly right-censored survival time. However, in many clinical trials, the true covariates may not always be accurately measured due to natural biological fluctuation or instrument error. It is well known that for regression analysis in general, naively using mismeasured covariates in conventional inference procedures may incur substantial estimation bias. In the presence of covariate measurement error, several functional modeling methods have been proposed under the situation where the distribution of the measurement error is known. Among them are parametric corrected score and conditional score. Although both methods are consistent, each suffers from severe problem of multiple roots or absence of appropriate root when the measurement error is substantial. The problem persists even when the sample size is practically large. We conduct a detailed investigation on the pathological behaviors of parametric corrected score and propose an approach of incorporating additional estimating functions to remedy these pathological behaviors. The estimation and inference are then accomplished by means of quadratic inference function. Extensive simulation studies are conducted to evaluate the performance of proposed method.

© 2015 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

The proportional hazards model is one of the most popular models to investigate the relationship between time to failure and covariates. However, in many clinical trials, the true covariates may not always be accurately measured due to natural biological fluctuation or instrument error. In some studies, the magnitude of measurement error could be substantial to the extent that it is comparable to or even larger than that of the true underlying covariate. A typical example is the HIV viral load in HIV/AIDS studies.

For regression analysis in general, naively using mismeasured covariates in conventional inference procedures may incur substantial estimation bias and several statistical methods have been suggested to address covariate measurement error; see the monograph of Carroll et al. for a good summary [1]. Regression calibration is used frequently to yield approximate estimation. However, it is well known that the regression calibration estimator is inconsistent in general. For consistent estimation, methods have been developed under either structural or functional modeling, i.e., with or without parametric distributional assumptions imposed on the true covariates. By definition, functional modeling approach might be more appealing, particularly for its robustness. Available functional modeling methods for Cox proportional hazards model include the conditional score [2], the parametric corrected score [3–5], and the nonparametric corrected score [5–8]. The idea of the conditional score is to condition away the nuisance parameters based on certain sufficient statistics whereas the last two classes adopt a correction strategy by constructing a corrected estimating function with error-contaminated covariates that shares the same limit as a reference estimating function with true underlying covariates. If the reference estimating function admits consistent estimates only, the corrected estimating function shall inherit this property in a compact parameter space containing the true value. Conditional score and parametric corrected score are generally different. But in the case of the Cox proportional hazards model and normal measurement error, the conditional score and parametric
corrected score estimators are asymptotically equivalent.

Although all three aforementioned methods produce consistent estimators, they all suffer from finite-sample pathological behaviors especially when the measurement error is substantial, which limit the applicability of these methods in practice. Recently, Huang [9] proposed an approach to incorporate additional estimating functions which constrain the derivatives of the parametric corrected score for loglinear model. This approach effectively remedies those pathological behaviors and also considerably improves the estimation efficiency. Huang’s approach provides a promising general strategy to handle similarly ill-behaved estimating functions.

In this paper, we first conduct a detailed investigation on pathological behaviors of parametric corrected score and conditional score. After that, we propose an augmented estimation procedure in which additional estimating functions are added to the parametric corrected score for the proportional hazards model and present the parametric corrected score for loglinear model. This approach effectively remedies those pathological behaviors and also considerably improves the estimation efficiency. Huang’s approach provides a promising general strategy to handle similarly ill-behaved estimating functions.

In this paper, we first conduct a detailed investigation on pathological behaviors of parametric corrected score and conditional score. After that, we propose an augmented estimation procedure in which additional estimating functions are added to the parametric corrected score for the proportional hazards model and present the parametric corrected score for loglinear model. This approach effectively remedies those pathological behaviors and also considerably improves the estimation efficiency. Huang’s approach provides a promising general strategy to handle similarly ill-behaved estimating functions.

### 2. Parametric corrected score and conditional score for proportional hazards model and their pathological behaviors

The proportional hazards model postulates that the cumulative hazard function \( \Lambda(\cdot) \) of survival time \( T \) of an individual with a \( p \)-vector of covariate \( Z \) has the form

\[
\Lambda(dt|Z) = \exp(\beta'Z)\Lambda_0(dt)
\]

where \( \beta \) is a \( p \)-vector parameter of interest and \( \Lambda_0(\cdot) \) is an unspecified underlying cumulative hazard function. Let \( C \) denote the censoring time and adopt the usual independent censoring mechanism: given \( Z, C \) is independent of \( T \).

The observed data, \( (X_i, \Delta, Z_i, i=1, \ldots, n) \), consist of \( n \) independent and identically distributed (i.i.d) replications of \( (X \sim F, \Delta \sim \mathcal{I}(T \leq C) | Z) \). The standard inference procedure for Cox proportional hazards model is then to maximize the partial likelihood or, equivalently, to solve the estimating function

\[
\xi^\ast(b, \hat{\Lambda}_0(\cdot)) = n^{-1} \sum_{i=1}^{n} \int_{0}^{T} \left( \frac{Z_i}{1 - \exp(b'Z_i)\Lambda_0(t)} \right) dN_i(t) - Y_i(t)\exp(b'Z_i)d\hat{\Lambda}_0(t),
\]

where \( N_i(t) = \int_{T \leq t \leq \Delta_i = 1} \) is the counting process, \( Y_i(t) = \int_{X_i \geq t} \) is the at-risk process and \( r \) is a positive constant such that \( Pr(T > r) > 0 \). Profiling out \( \Lambda_0(\cdot) \), the estimating function for \( \beta \) alone is

\[
\xi(b) = n^{-1} \sum_{i=1}^{n} \int_{0}^{T} \left( Z_i - \frac{\sum_{j=1}^{n} Y_j(t) Z_j \exp(b'Z_j)}{\sum_{j=1}^{n} Y_j(t) \exp(b'Z_j)} \right) dN_i(t).
\]

Estimating function (2) is actually the usual partial score function.

#### 2.1. Parametric corrected score and conditional score

Split covariates \( Z = [Z_d', Z_o']^T \) where \( Z_o \) are those covariates that can be accurately measured and \( Z_d \) are covariates prone to measurement error and cannot be accurately measured. Though \( Z_o \) cannot be measured directly, we can observe them through their surrogates \( W \).

Under the classical additive measurement error model, \( W = Z_o + e \), where \( e_r \) is the error vector and \( e_r \) is assumed to be independent of \( (\Delta, Z) \). In this paper, we assume that the distribution of \( e_r \) is known; The situation where distribution of \( e_r \) is unknown will be discussed in Section 5.

Let \( W = (Z_d, W_o) \) and \( e = (e_r, e_o) \). The observed data now consist of \((X_i, \Delta, W_i, i=1, \ldots, n)\) in the presence of covariate measurement error. It is well known that naively replacing \( Z \) by \( W \) in estimating functions (1) or (2) could incur substantial estimation bias. Denote the cumulant-generating function of \( e \) as \( \Omega(b) = \log \mathcal{E}(\exp(b'\epsilon)) \) and its derivative \( \Omega'_r(b) = \Omega'(b)/\partial b \). The parametric corrected score estimating function is given by

\[
\eta^\ast(b, \hat{\Lambda}_0(\cdot)) = n^{-1} \sum_{i=1}^{n} \int_{0}^{T} \left( W_i - \hat{\Omega}(0) \right) dN_i(t) - \frac{\sum_{j=1}^{n} Y_j(t) W_j \exp(b'W_j)}{\sum_{j=1}^{n} Y_j(t) \exp(b'W_j)} \right) dN_i(t).
\]

Further profiling out \( \hat{\Lambda}_0(\cdot) \), we obtain the corrected score for \( \beta \):

\[
\eta(b) = n^{-1} \sum_{i=1}^{n} \int_{0}^{T} \left( W_i + \hat{\Omega}(b) - \hat{\Omega}(0) \right) dN_i(t)
\]

which has the same expectation as reference (2) asymptotically for each and every finite \( b \). The estimation is then to find the zero crossing of the above estimating function. The consistency and asymptotic normality of corrected score estimator are later established by Kong and Gu [11].

Huang and Wang [6] defined a root-consistent estimating function as such that every zero-crossing is consistent and showed that a normalized estimating function is root-consistent if its limit has a unique root at the estimand. By definition, reference (2) is a root-consistent estimating function and the new estimating function (4) shall inherit the root-consistency from (2). The root-consistency of (4) assures that in a compact parameter space containing the true parameter \( \beta \), the parametric corrected score will admit a unique root asymptotically and the root is consistent and asymptotically normal.

When the measurement error is normally distributed and assume the variance matrix is \( \Sigma \), the conditional score estimating function for the Cox proportional hazards model [12] can be written as

\[
\eta_{con}(b) = n^{-1} \sum_{i=1}^{n} \int_{0}^{T} \left( W_i + \Sigma b \right) dN_i(t)
\]

In fact, it can be shown that estimators from parametric
corrected score and conditional score are asymptotically equivalent in the case of normal measurement error.

2.2. Pathological behaviors

When the magnitude of measurement error is small, the asymptotic results of parametric corrected score and conditional score provide a good approximation for practical purposes. However, when the measurement error increases, the pathological behaviors may start to arise [13]. These pathological behaviors include multiple zero-crossings or a single wrong zero-crossing that is inappropriate. These pathological behaviors may cause serious concerns when the measurement error is substantial and limit the applicability of parametric corrected score and conditional score in practice. In this section, we will conduct a detailed investigation of pathological behaviors for these two methods.

We first consider the parametric corrected score. Consider a single-covariate model with normal measurement error. In the absence of measurement error, the partial likelihood score function \( \hat{\gamma}(b) \) is monotonically decreasing and has a unique root. The asymptotic result suggests that the parametric corrected score \( h(b) \) should be monotonically decreasing as well in a compact parameter space containing the true parameter when the sample size is large. One may speculate the parametric corrected score to have an overall decreasing trend over the entire parameter space. But surprisingly, the overall trend of parametric corrected score in an unbounded parameter space is increasing. Function \( h(b) \) takes a value of \(-\infty\) when \( b = -\infty \) and a value of \(+\infty\) when \( b = +\infty \). This observation suggests that the parametric corrected score \( h(b) \) has an odd number of roots. In our numerical studies, only single- and triple-root patterns have been observed and two typical plots of parametric corrected score are illustrated in Fig. 1. The same root patterns were observed in Huang’s investigation of loglinear model [9].

If we characterize a root by increase or decrease of \( h(b) \) around it, the increasing and decreasing roots correspond to local minimizers and maximizers of corresponding objective function, respectively. Therefore, an increasing root is considered as an inappropriate one. In the case of single-root pattern, the only root is increasing and thus inappropriate. For the triple-root pattern, an appropriate root exists since there is only one decreasing root. The single-root pattern is considered as root-finding failure.

We conduct a simulation study to examine the prevalence of single-root pattern. We consider a single covariate Z and generate it from various distributions: A) standard normal distribution, B) modified chi-square distribution, and C) uniform distribution with mean 0 and variance 1. To generate the modified chi-square distribution, the chi-square distribution with 1 degree of freedom was first truncated at 5 and then location-shifted to mean 0 and rescaled to variance 1. The measurement error follows standard normal distribution. The true coefficient was taken to be \( \beta = -1 \) and the baseline hazard is constant 1. Censoring was generated from a uniform distribution on \([0,\mu]\), where \( \mu \) is chosen so that the censoring rate is ranged from 20% to 60%. These set-ups represent a practical scenario with substantial error contamination on the covariate. The results based on 1000 iterations are reported in Table 1. The prevalence of single-root pattern is similar across different scenarios with various censoring rates. When the sample size is 100, the percentage of single-root pattern under all three distributions are close to or over 60%. Even when the sample size increases to 800, the prevalence of single-root pattern is still quite

Table 1

<table>
<thead>
<tr>
<th>Censoring rate</th>
<th>Distribution of X</th>
<th>Size</th>
<th>100</th>
<th>200</th>
<th>400</th>
<th>800</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>Normal</td>
<td>68.0</td>
<td>52.1</td>
<td>38.0</td>
<td>20.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Modified chi-square</td>
<td>65.8</td>
<td>57.9</td>
<td>52.3</td>
<td>40.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>64.5</td>
<td>51.9</td>
<td>37.6</td>
<td>19.6</td>
<td></td>
</tr>
<tr>
<td>40%</td>
<td>Normal</td>
<td>62.4</td>
<td>49.0</td>
<td>36.3</td>
<td>17.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Modified chi-square</td>
<td>64.4</td>
<td>58.0</td>
<td>50.9</td>
<td>43.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>60.4</td>
<td>49.5</td>
<td>37.1</td>
<td>19.3</td>
<td></td>
</tr>
<tr>
<td>60%</td>
<td>Normal</td>
<td>61.0</td>
<td>46.5</td>
<td>32.1</td>
<td>16.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Modified chi-square</td>
<td>62.7</td>
<td>57.3</td>
<td>50.9</td>
<td>42.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uniform</td>
<td>58.5</td>
<td>47.9</td>
<td>36.8</td>
<td>21.2</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Observed root patterns of the parametric corrected score \( h(b) \). The true \( \beta \) is \(-1\) and these two corrected curves correspond to the same profile score (with true covariates). Portion of a corrected curve is thickened to indicate negative derivative.
For the conditional score, the patterns are more complicated. When the absolute value of $b$ gets large, the estimating function fluctuates around zero and finally approaches zero as $b$ goes to infinity. Therefore the conditional score may have many zero-crossings. When the $b$ is not so extreme, two general patterns for the conditional scores are observed and plots from two simulation datasets are shown in Fig. 2. In the first pattern, the conditional score has one zero-crossing close to the true parameter. This single zero-crossing is decreasing and thus an appropriate one. In the second pattern, the conditional score appears to have no proper zero-crossing near the truth though it may have multiple zero-crossings at extreme values of $b$. For conditional score, we define root-finding failure as following: we start with the naive estimator and search locally in the direction of the derivative of estimating function for a new estimator such that the $\| \cdot \|^2$-norm of the estimating functions decreases after each step. Repeat this step and we will eventually achieve a root or a local minimizer for the $\| \cdot \|^2$-norm of the estimating functions. We consider the latter case as root-finding failure. Table 2 summarizes the prevalence of root-finding failure for the conditional score. The same simulation set up as in the corrected score is used. With a sample size of 100, the root-finding failure rate for conditional score varies from 3% to 5% for normal covariate and from 14% to 25% for modified chi-square covariate. As the sample size increases to 800, the failure rate drops to 0.2% for normal covariate but remains at least 14% for modified chi-square covariate.

The above investigation results show that, in the presence of substantial measurement error, both parametric corrected score and conditional score suffer from severe finite-sample pathological behaviors. Therefore improvements are required for these methods to have practical applicability. We observe that an appropriate zero-crossing of an estimating function should be a decreasing one. This observation suggests that the trend of estimating function is also informative and could be taken into account in the estimate determination. By Taylor expansion, the trend of estimating function may be quantified by its derivative. Recognizing this feature, Huang [9] proposed an approach to incorporate additional estimating functions which constrain the derivatives of the corrected score for the loglinear model. The estimation and inference are then accomplished by means of empirical likelihood. This approach effectively remedies the pathological behaviors of corrected score for loglinear model and also considerably improves the estimation efficiency. However, in the case of the Cox proportional hazards model, we are unable to construct additional estimating functions that effectively constrain the derivative of the parametric corrected score or conditional score because of the very nature of these two estimating functions. Nevertheless, Huang’s approach provides an insight into a new approach to address pathological behaviors of estimating functions. If we could identify additional estimating functions that do not share the same wrong roots as the original estimating function, then by combining the original and additional estimating functions, pathological behaviors could be reduced or eliminated. But if either the original estimating function or the additional estimating functions vanish to zero, then wrong root sharing would easily arise. We have shown in the simulation that the conditional score would vanish to zero when the absolute value of parameter becomes large. Thus, the trend pattern of the parametric corrected score is more desirable than that of the conditional score and our method will be developed for the parametric corrected score.

### 3. Improving corrected score

#### 3.1. Augmented estimation method

Motivated by Huang’s trend-constrained corrected score [9], we
first derive the following result:

**Theorem 1.** Under the proportional hazards model and the classical additive measurement error model,

$$
\frac{\tau}{\tau} \int_{0}^{\tau} \frac{\partial^{k_{i}+\cdots+k_{p}}}{\partial b_{i}^{k_{i}} \cdots b_{p}^{k_{p}}} \exp \{ b'W - \Omega(b) \} dN_{t}(t) \Big|_{b=0} - \Omega(0) \int_{0}^{\tau} Y(t) \exp \{ b'W - \Omega(b) \} d\Lambda_{0}(t) \Big|_{b=0} = 0.
$$

for $k_{l} \geq 0, l = 1, \ldots, p$, where $b_{l}, l = 1, \ldots, p$, is the $l$th element of $b$.

Equation (6) is useful in constructing additional estimating equations for $\hat{\beta}$. When $\sum_{l=1}^{p} k_{l} = 0$ and 1, one may obtain the usual parametric corrected score. When $\sum_{l=1}^{p} k_{l} = 2$, the additional estimating functions are the upper triangular elements of the following symmetric matrix:

$$
\frac{\tau}{\tau} \int_{0}^{\tau} \frac{\partial^{k_{i}+\cdots+k_{p}}}{\partial b_{i}^{k_{i}} \cdots b_{p}^{k_{p}}} \exp \{ b'W_i - \Omega(b) \} dN_{t}(t) - \frac{\tau}{\tau} \int_{0}^{\tau} Y_i(t) \exp \{ b'W_i - \Omega(b) \} d\Lambda_{0}(t).
$$

By profiling out $\Lambda_{0}(\cdot)$, we obtain

$$
\frac{\tau}{\tau} \int_{0}^{\tau} \frac{\partial^{k_{i}+\cdots+k_{p}}}{\partial b_{i}^{k_{i}} \cdots b_{p}^{k_{p}}} \exp \{ b'W_i - \Omega(b) \} - \Omega(0) \int_{0}^{\tau} Y_i(t) \exp \{ b'W_i - \Omega(b) \} d\Lambda_{0}(t).
$$

The additional estimating functions would be helpful if both parametric corrected score and additional estimating function are close to 0 around the truth (not necessarily having roots) and the additional estimating function is not close to 0 when the parameter of interest is not close to each other. This is an ideal scenario. The parametric corrected score have three zero-crossings. The additional estimating function shares the same decreasing zero-crossing as the parametric corrected score. Moreover, they do not share any wrong zero-crossings. In plot (a), the parametric corrected score have three zero-crossings and two of them are close to each other. In this case, when the two roots around the truth are very close to 0, the additional estimating function is not close to 0 near this wrong zero-crossing. Both estimating functions are close to 0 around the truth.

### 3.2. Estimation and inference

With the additional estimating functions, we have more estimating functions than the number of parameters. Available methods to synthesize estimating functions that exceed the number of parameters include empirical likelihood [14] and quadratic inference function (QIF) method [15]. In this research, we shall use the quadratic inference function method to determine the estimate since the estimating functions are not sums of iid terms, thus the empirical likelihood would be computational difficult.

Let $\varphi(b)$ denotes the estimating functions. $\varphi(b)$ is comprised of the original parametric corrected score and additional estimating functions. The quadratic inference function takes the form

$$
Q(b; \hat{b}) = \varphi'(b) \hat{C}^{-1}(b) \varphi(b),
$$

where $\hat{C}(b)$ is any consistent estimator for the asymptotic variance of $\varphi(b)$. Then the estimator is defined as the minimizer of (8) and is consistent for the true value of $b$. Furthermore, the estimator is asymptotically efficient in the class of consistent estimators based on the quadratic inference function. We name the method to incorporate additional estimating functions based on (6) as the augmented parametric corrected score. One important special case is the method incorporating the upper triangular elements of matrix (7) and was termed the second order augmented parametric corrected score. Augmented parametric corrected scores with higher order are also available, with additional estimating functions corresponding to $k_{l}$ such that $\sum_{l=1}^{p} k_{l} > 2$. In Section 4, we will conduct extensive simulation studies to evaluate the performance of the augmented parametric corrected score and its applicability in practice.

### 4. Numerical studies

#### 4.1. Simulations

Extensive simulations studies were conducted to evaluate the performance of augmented parametric corrected score. For reference and comparison, the ideal, naive, regression calibration, conditional score, and parametric corrected score were also studied.
The ideal estimator used the ordinary partial score function with true covariates and, of course, it is not a realistic estimator. The naive approach uses the mismeasured surrogates in place of true covariates in the partial score function. For the regression calibration method, $Z_e$ is replaced by $E(Z_e|Z_a, W_e)$ in the partial score function. For the proposed approach, the optimization algorithm of Nelder and Mead\cite{16} will be used.

As shown in previous section, both conditional score and parametric corrected score have high prevalence of root-finding failure. To utilize all simulated data and conduct a fair comparison, we propose the following re-defined conditional score and re-defined parametric corrected score. If an appropriate zero-crossing could be found, the estimators will take the value of zero-crossing. But if root-finding failure occurs, the estimators will be defined as the local minimizer of the $2$-norm of the estimating functions closest to the naive estimator. Operationally, we will use the following modified Newton–Raphson algorithm. We start with the naive estimator and calculate the Newton–Raphson step size. Since the goal is to find a root or local minimizer, we need prevent overshooting. In our simulation, we cap the step size at $0.2$. During each iteration, we compare the $2$-norm evaluated at new estimator to that evaluated at current estimator. If the $2$-norm evaluated at new estimator is smaller, then the new estimator will be accepted and the algorithm continues to the next iteration. Otherwise, we will halve the step size and calculate the new estimator again. Iterations will be repeated until that i) the absolute value of estimating function is less than $10^{-6}$ or ii) the step size has been halved for more than 10 time during any single iteration. If criteria i) is satisfied, the algorithm converges to a zero-crossing and a root is identified in this case. If criteria ii) is satisfied, the algorithm converges to a local minimizer and root-finding failure occurs. Comparing to the original definition that estimators take the values of zero-crossings of estimating function, this new definition actually benefits the conditional score and corrected score. As shown in Figs. 1 and 2, in the case of root-finding failure, zero-crossings of conditional score and corrected score are at extreme and far away from the true value.

In the simulation, we consider both single- and double-covariate models. In the single-covariate models, the true covariate $X$ is of mean 0 and variance 1. The true regression coefficient was set to be $-1$ and the baseline hazard is constant 1. The measurement error follows the standard normal distribution. Two different distributions of $X$ were studied: A) standard normal distribution and B) modified chi-square distribution. To generate the modified chi-square distribution, the chi-square distribution with 1 degree of freedom was first truncated at 5 and then location-shifted.

![Parametric corrected score and additional estimation function based on Equation (7) for a single-covariate model with $\beta = -1$.](image)

---

Fig. 3. Parametric corrected score and additional estimation function based on Equation (7) for a single-covariate model with $\beta = -1$. 
to mean 0 and rescaled to variance 1. Censoring time was generated from a uniform distribution on [0, $\mu$] and we will consider two different settings of censoring rate at 20% and 60%.

In the double-covariate models, true covariate $X$ follows bivariate normal distribution with mean (0,0), variance (1,1) and correlation coefficient of .5. The first covariate was subject to a standard normal measurement error, whereas the second covariate was accurately measured. The regression coefficients were set to ($C_0$, $C_1$), and the baseline hazard is constant 1. Censoring time was also generated from a uniform distribution on [0, $m$], where $m$ is chosen so that the censoring rate is 20% or 60%.

Sample sizes 100, 200, 400, 800, and 1600 were investigated. For each scenario, 1000 samples were simulated. We report the results on point and interval estimation separately.

Tables 3 and 4 summarize the simulation results on the estimators in the single-covariate models with censoring rates of 20% and 60% respectively. The quantile-quantile plots are shown in Figs. 5 and 6. For each scenario, the mean bias, and standard deviation were calculated. For augmented parametric corrected score, three sets of additional estimating functions were considered where $k = 2, 3, 4$. As expected, the naive estimator has substantial bias under both scenarios. The regression calibration estimator shows moderate bias, with larger bias in modified chi-square covariate case than normal covariate case. The re-defined conditional score shows slight bias under both scenarios, probably due to its left skewness. The quantile-quantile plots show that the re-defined conditional score deviates from normality considerably even when the sample size is 1600. It also has a much larger standard deviation comparing to the re-defined parametric corrected score and the augmented parametric corrected score. The re-defined parametric corrected score is unbiased for both scenarios and the standard deviation is small. All three augmented corrected scores are consistent under both scenarios, and they become less biased as sample size increases. In comparison with higher-order augmented corrected scores, the second order augmented corrected score seems more favorable overall in terms of bias and standard error. Compared to the re-defined parametric corrected score, the second order augmented corrected score has a larger

![Fig. 4. Quadratic inference functions for the two datasets in Fig. 1.](image)

### Table 3

Simulation summary statistics for the single-covariate models with 20% censoring rate: Ideal, naive (NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and first k augmented parametric corrected score (ACS: k), $k = 2, 3, 4$.  

<table>
<thead>
<tr>
<th>Size</th>
<th>Ideal</th>
<th>NV</th>
<th>RC</th>
<th>ConS</th>
<th>CS</th>
<th>ACS: 2</th>
<th>ACS: 3</th>
<th>ACS: 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
<td>F</td>
<td>B</td>
<td>SD</td>
<td>F</td>
</tr>
<tr>
<td>Normal covariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>−15</td>
<td>154</td>
<td>593</td>
<td>99</td>
<td>148</td>
<td>254</td>
<td>2.7</td>
<td>−376</td>
</tr>
<tr>
<td>200</td>
<td>−8</td>
<td>104</td>
<td>602</td>
<td>65</td>
<td>189</td>
<td>163</td>
<td>1.5</td>
<td>−251</td>
</tr>
<tr>
<td>400</td>
<td>−5</td>
<td>73</td>
<td>601</td>
<td>48</td>
<td>195</td>
<td>113</td>
<td>.6</td>
<td>−202</td>
</tr>
<tr>
<td>800</td>
<td>1</td>
<td>49</td>
<td>604</td>
<td>32</td>
<td>206</td>
<td>76</td>
<td>.3</td>
<td>−99</td>
</tr>
<tr>
<td>1600</td>
<td>1</td>
<td>36</td>
<td>605</td>
<td>23</td>
<td>208</td>
<td>53</td>
<td>0</td>
<td>−43</td>
</tr>
<tr>
<td>$\chi^2$ Covariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>−25</td>
<td>204</td>
<td>639</td>
<td>91</td>
<td>242</td>
<td>247</td>
<td>14.4</td>
<td>−485</td>
</tr>
<tr>
<td>200</td>
<td>−5</td>
<td>136</td>
<td>640</td>
<td>63</td>
<td>267</td>
<td>155</td>
<td>12.2</td>
<td>−396</td>
</tr>
<tr>
<td>400</td>
<td>−6</td>
<td>98</td>
<td>642</td>
<td>43</td>
<td>280</td>
<td>101</td>
<td>15.1</td>
<td>−329</td>
</tr>
<tr>
<td>800</td>
<td>−3</td>
<td>70</td>
<td>641</td>
<td>32</td>
<td>281</td>
<td>72</td>
<td>13.9</td>
<td>−249</td>
</tr>
<tr>
<td>1600</td>
<td>0</td>
<td>47</td>
<td>644</td>
<td>21</td>
<td>289</td>
<td>47</td>
<td>12.5</td>
<td>−139</td>
</tr>
</tbody>
</table>

Note: F: root-finding failure (%); B: mean bias (×10$^3$); SD: standard deviation (×10$^3$).
standard deviation when the sample size is small. But the standard
deviation of augmented corrected score decreases rapidly as the
sample size increases and is smaller than that of the re-defined
corrected score when the sample size is 1600.

Tables 5 and 6 show the simulation results for the double-
covariate models. As expected, when two covariates are corre-
lated, the measurement error generally has impact not only on the
mismeasured covariate but also on that of the accurately measured
one. The relative performance of all estimators is similar to what
was observed in the single-covariate models. For multiple-
covariate models, each set of additional estimating functions con-
tains more than one element. For example, when $k = 2$, the addi-
tional estimating functions includes three elements in the upper
triangle of matrix (3.7). Various augmented corrected score could
be constructed depending on the additional estimating functions
chosen. Since the second order augmented corrected score shows a

Table 4
Simulation summary statistics for the single-covariate models with 60% censoring rate: Ideal, naive (NV), regression calibration (RC), re-defined conditional score (ConS), re-defined parametric corrected score (CS), and first $k$ augmented parametric corrected score (ACS: $k$), $k = 2, 3, 4$.

<table>
<thead>
<tr>
<th>Size</th>
<th>Ideal</th>
<th>NV</th>
<th>RC</th>
<th>ConS</th>
<th>CS</th>
<th>ACS: 2</th>
<th>ACS: 3</th>
<th>ACS: 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
</tr>
<tr>
<td>Normal covariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>24</td>
<td>206</td>
<td>550</td>
<td>134</td>
<td>57</td>
<td>335</td>
<td>4.7</td>
<td>-429</td>
</tr>
<tr>
<td>200</td>
<td>15</td>
<td>142</td>
<td>562</td>
<td>88</td>
<td>109</td>
<td>206</td>
<td>2.5</td>
<td>-272</td>
</tr>
<tr>
<td>400</td>
<td>8</td>
<td>97</td>
<td>563</td>
<td>63</td>
<td>119</td>
<td>142</td>
<td>.7</td>
<td>-197</td>
</tr>
<tr>
<td>800</td>
<td>2</td>
<td>65</td>
<td>567</td>
<td>42</td>
<td>132</td>
<td>96</td>
<td>.3</td>
<td>-100</td>
</tr>
<tr>
<td>1600</td>
<td>1</td>
<td>46</td>
<td>567</td>
<td>30</td>
<td>133</td>
<td>69</td>
<td>0</td>
<td>-51</td>
</tr>
<tr>
<td>$\chi^2$ covariate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>-66</td>
<td>351</td>
<td>687</td>
<td>124</td>
<td>351</td>
<td>288</td>
<td>24.6</td>
<td>-259</td>
</tr>
<tr>
<td>200</td>
<td>-24</td>
<td>239</td>
<td>691</td>
<td>84</td>
<td>370</td>
<td>190</td>
<td>21.7</td>
<td>-291</td>
</tr>
<tr>
<td>400</td>
<td>-16</td>
<td>159</td>
<td>693</td>
<td>57</td>
<td>382</td>
<td>124</td>
<td>25.7</td>
<td>-202</td>
</tr>
<tr>
<td>800</td>
<td>-9</td>
<td>112</td>
<td>693</td>
<td>40</td>
<td>385</td>
<td>86</td>
<td>25.0</td>
<td>-172</td>
</tr>
<tr>
<td>1600</td>
<td>0</td>
<td>78</td>
<td>697</td>
<td>28</td>
<td>394</td>
<td>60</td>
<td>21.5</td>
<td>-77</td>
</tr>
</tbody>
</table>

Note: Same as in that of Table 3.

Fig. 5. Quantile-quantile plots for $\beta$ in the single-covariate models with 20% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1600.
more favorable overall performance in the single-covariates models, we consider only the second order augmented corrected score and augmented corrected scores using a subset of the three elements in this simulation. Simulation results show that the second order augmented corrected score performs better than two other augmented corrected scores. The bias of the second order augmented corrected score reduced quickly as sample size increased. Meanwhile, it also has the smallest standard deviation.

Tables 7 and 8 report the coverage of three types of 95% confidence intervals in the single- and double-covariate models. All three scenarios use the second order augmented corrected score. In constructing the confidence interval, we use two different approaches: inverting the hypothesis testing statistics as introduced in Section 2.2, and the Wald-type confidence interval. For the former, we use two critical values based on the asymptotic chi-square distribution and the bootstrap calibration [17]. Bootstrap size of

![Fig. 6. Quantile-quantile plots for $\beta$ in the single-covariate models with 60% censoring rate, where $\beta = -1$. Red, yellow, green, blue, and black correspond to sample sizes 100, 200, 400, 800, and 1600.](image)
Note: Same as in that of Table 5.

Table 7

Coverage of 95% confidence interval for the second order augmented parametric corrected score with 20% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval.

<table>
<thead>
<tr>
<th>Size</th>
<th>Ideal</th>
<th>NV</th>
<th>BC</th>
<th>ConS</th>
<th>CS</th>
<th>ACS: 1/3</th>
<th>ACS: 2/3</th>
<th>ACS: 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
</tr>
<tr>
<td>100</td>
<td>b₁</td>
<td>32</td>
<td>242</td>
<td>610</td>
<td>140</td>
<td>-13</td>
<td>513</td>
<td>43.3</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>33</td>
<td>245</td>
<td>-352</td>
<td>206</td>
<td>-40</td>
<td>336</td>
<td>-27</td>
</tr>
<tr>
<td>200</td>
<td>b₁</td>
<td>-20</td>
<td>160</td>
<td>617</td>
<td>95</td>
<td>59</td>
<td>286</td>
<td>36.1</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>21</td>
<td>161</td>
<td>-362</td>
<td>138</td>
<td>-81</td>
<td>209</td>
<td>24</td>
</tr>
<tr>
<td>400</td>
<td>b₁</td>
<td>-6</td>
<td>109</td>
<td>619</td>
<td>65</td>
<td>97</td>
<td>182</td>
<td>21.5</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>6</td>
<td>111</td>
<td>-366</td>
<td>97</td>
<td>-106</td>
<td>137</td>
<td>31</td>
</tr>
<tr>
<td>800</td>
<td>b₁</td>
<td>-1</td>
<td>72</td>
<td>621</td>
<td>45</td>
<td>112</td>
<td>120</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>2</td>
<td>76</td>
<td>-369</td>
<td>69</td>
<td>-115</td>
<td>94</td>
<td>34</td>
</tr>
<tr>
<td>1600</td>
<td>b₁</td>
<td>-2</td>
<td>53</td>
<td>622</td>
<td>33</td>
<td>114</td>
<td>87</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>0</td>
<td>53</td>
<td>-371</td>
<td>46</td>
<td>-117</td>
<td>65</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 8

Coverage of 95% confidence interval for the second order augmented parametric corrected score with 60% censoring rate. C (chi-square distribution), BC (bootstrap calibration) and W (Wald-type) indicate the type of confidence interval.

<table>
<thead>
<tr>
<th>Size</th>
<th>Ideal</th>
<th>NV</th>
<th>BC</th>
<th>ConS</th>
<th>CS</th>
<th>ACS: 1/3</th>
<th>ACS: 2/3</th>
<th>ACS: 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
<td>B</td>
<td>SD</td>
</tr>
<tr>
<td>100</td>
<td>b₁</td>
<td>32</td>
<td>242</td>
<td>610</td>
<td>140</td>
<td>-13</td>
<td>513</td>
<td>43.3</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>33</td>
<td>245</td>
<td>-352</td>
<td>206</td>
<td>-40</td>
<td>336</td>
<td>-27</td>
</tr>
<tr>
<td>200</td>
<td>b₁</td>
<td>-20</td>
<td>160</td>
<td>617</td>
<td>95</td>
<td>59</td>
<td>286</td>
<td>36.1</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>21</td>
<td>161</td>
<td>-362</td>
<td>138</td>
<td>-81</td>
<td>209</td>
<td>24</td>
</tr>
<tr>
<td>400</td>
<td>b₁</td>
<td>-6</td>
<td>109</td>
<td>619</td>
<td>65</td>
<td>97</td>
<td>182</td>
<td>21.5</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>6</td>
<td>111</td>
<td>-366</td>
<td>97</td>
<td>-106</td>
<td>137</td>
<td>31</td>
</tr>
<tr>
<td>800</td>
<td>b₁</td>
<td>-1</td>
<td>72</td>
<td>621</td>
<td>45</td>
<td>112</td>
<td>120</td>
<td>10.8</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>2</td>
<td>76</td>
<td>-369</td>
<td>69</td>
<td>-115</td>
<td>94</td>
<td>34</td>
</tr>
<tr>
<td>1600</td>
<td>b₁</td>
<td>-2</td>
<td>53</td>
<td>622</td>
<td>33</td>
<td>114</td>
<td>87</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td>b₂</td>
<td>0</td>
<td>53</td>
<td>-371</td>
<td>46</td>
<td>-117</td>
<td>65</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 9

Comparison of regression coefficient estimators in the ACTG 175 data.

<table>
<thead>
<tr>
<th>log(CD4)</th>
<th>ZDV + ddl</th>
<th>ZDV + ddC</th>
<th>ddl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Est</td>
<td>Var</td>
<td>Est</td>
<td>Var</td>
</tr>
<tr>
<td>NV</td>
<td>-1.839</td>
<td>-1.183</td>
<td>0.881</td>
</tr>
<tr>
<td>CS</td>
<td>-2.177</td>
<td>1.678</td>
<td>0.991</td>
</tr>
<tr>
<td>ACS</td>
<td>-1.785</td>
<td>1.422</td>
<td>0.991</td>
</tr>
</tbody>
</table>

Note: For proposed estimator, the first row of variance estimator is obtained by inverting hypothe single testing statistics with bootstrap critical value; the second row of values is from sandwich variance estimator. Est: Estimated coefficient; Var: Variance.
We consider a Cox regression model with 4 covariates: the true baseline log(CD4) and three indicators for the four treatments with ZDV group as the reference. We define the baseline log(CD4) as the average of the two log(CD4) measurements. From the duplicated measurements, we estimated the variance for error and true underlying log(CD4) to be .033 and .076 respectively. Note that the variance of measurement error is estimated using two replicated measurements of baseline log(CD). Therefore there is an additional underlying log(CD4) to be .033 and .076 respectively. Note that the measurements, we estimated the variance for error and true un-
average of the two log(CD4) measurements. From the duplicated baseline log(CD4) and three indicators for the four treatments with 
carcinoma (BCC) and squamous cell carcinoma (SCC) [18]. Baseline 
trial, which involved multiple recurrences of skin cancers, basal cell 
carcinoma (BCC) and squamous cell carcinoma (SCC) [18]. Baseline 
measurement tool is another contributor. One example of mea-

5. Discussion

Measurement error is a common issue for many clinical trials. The main reason for the presence of measurement error is study participants’ natural biological fluctuation. The imprecision of measurement tool is another contributor. One example of measurement error is from the Nutritional Prevention of Cancer (NPC) measurement tool is another contributor. One example of measurement error in clinical trial.

For proportional hazards model with covariate measurement error, several consistent methods have been proposed under the functional modeling framework, including the conditional score and the parametric corrected score. However, when the measurement error is substantial to the extent that the errors are comparable to the true covariates in variance, both methods might experience pathological behaviors and root finding failure. Recently, Huang [9] developed a novel approach to incorporate additional estimating functions which constrain the derivatives of the parametric corrected score. That approach proves effective and eliminates finite sample pathological behaviors of parametric corrected score for the loglinear model. Motivated by Huang’s approach, we conduct an investigation on the pathological behaviors of parametric corrected score and conditional score and propose an augmented parametric corrected score for the proportional hazards model by incorporating additional estimating functions to the original parametric corrected score. Results of simulation studies show the proposed approach is effective in eliminating pathological behaviors even with small sample size and substantial measurement error. The variance of proposed estimator appears to be larger than the parametric corrected score when sample size is smaller than 400, but it decreases rapidly and become smaller than the parametric corrected score as the sample size increases to 400. With the ability of eliminating finite sample pathological behaviors of conditional score and the parametric corrected score, this proposed augmented parametric corrected score could be widely used for time-to-event data in clinical trials where covariate measurement error issue is of concern.

In this paper, we have only considered the situation where the distribution of the measurement error is known. With additional data available on the measurement error, the parametric distribution imposed on the measurement error may be spared. With the availability of replicated mismeasured covariates, Huang and Wang [7] developed a nonparametric corrected score method for the proportional hazards model. Extension of the approach of incorporating additional estimating functions to the nonparametric corrected score is currently under study.

Appendix

A Asymptotic variance of estimating function

For simplicity, we consider the single-covariate model given in Section 2.2. The second order augmented corrected score has the form

$$
\varphi(b) = n^{-1} \sum_{i=1}^{n} \left\{ \left( \frac{W_i + b \sigma^2}{W_i^2 - \sigma^2} \right) - \frac{\sum_{j=1}^{n} Y_j(t) \exp(bW_j) \left( \frac{W_j}{(W_j - b \sigma^2)^2 - \sigma^2} \right)}{\sum_{j=1}^{n} Y_j(t) \exp(bW_j)} \right\} dN_i(t).
$$

With functional delta method [7], straightforward algebra gives

$$
n^{1/2} \varphi(b) = n^{-1/2} \sum_{i=1}^{n} (B_{11} - B_{12}) + o_p(1)
$$

where

$$
B_{11} = \int_0^\infty \left\{ \frac{\exp(bW_i) \left( \frac{W_i}{(W_i - b \sigma^2)^2 - \sigma^2} \right)}{\exp(bW_i)} \right\} dN_i(t), \quad \text{and}
$$

$$
B_{12} = \int_0^\infty \left\{ \frac{\exp(bW_i) \left( \frac{W_i}{(W_i - b \sigma^2)^2 - \sigma^2} \right)}{\exp(bW_i)} \right\} \frac{Y_i(t) dN(t)}{\exp(bW_i)}
$$

Thus, $n^{1/2} \varphi(b)$ is asymptotically a sum of iid random variables. For fixed $b$, $n^{1/2} \varphi(b)$ is asymptotically normal with a covariance matrix $\Sigma(b)$ that can be consistently estimated by

$$
\hat{\Sigma}(b) = n \sum_{i=1}^{n} \left( \bar{o}_i(b) - \bar{\tau}(b) \right) \left( \bar{o}_i(b) - \bar{\tau}(b) \right)^t
$$

where $\bar{o}_i(b) = n^{-1}(B_{i3} - B_{i4})$, $\bar{\tau}(b) = n^{-1} \sum_{i=1}^{n} \bar{o}_i(b)$, and $B_{i3}$ and $B_{i4}$ are defined as
\[ B_{13} = \int_{0}^{\tau} \left\{ \left( \frac{W_i + b \sigma^2}{W_i - \sigma^2} \right) - \frac{\tau(Y(t)\exp(bW))}{\hat{\epsilon}(Y(t)\exp(bW))} \right\} dN_i(t) \] and

\[ B_{14} = \int_{0}^{\tau} \left\{ \left( \frac{\exp(bW_i)}{(W_i - b \sigma^2)^2 - \sigma^2} \right) - \frac{\hat{\epsilon}(Y(t)\exp(bW))}{[\hat{\epsilon}(Y(t)\exp(bW))]^2} \right\} \exp(bW_i) \right\} \exp(bW_i) \left\{ \left( \frac{W}{(W - b \sigma^2)^2 - \sigma^2} \right) \right\} \exp(bW_i) \right\} \] \[ Y_i(t) d\hat{\epsilon}N(t). \]

Asymptotic variance of other augmented corrected scores could be derived similarly.

**B Proof of Theorem 1**

Given

\[ \epsilon(\exp(b'W - \Omega(b))|Z) = \exp(b'Z) \]

under additive measurement error model, one may obtain

\[ \epsilon \left[ \int_{0}^{\tau} Y \frac{\exp(b'W - \Omega(b))}{\partial b_k \cdots b_p} d\Lambda_0(t) \right] Z = \prod_{l=1}^{p} \int_{0}^{\tau} Y(t) \exp(b'Z) d\Lambda_0(t) \]

and

\[ \epsilon \left[ \int_{0}^{\tau} \frac{\partial^{k+\cdots+k} \exp(b'W - \Omega(b))}{\partial b_k \cdots b_p} d\Lambda_0(t) \right] Z = \prod_{l=1}^{p} \int_{0}^{\tau} Y(t) \exp(b'Z) d\Lambda_0(t) \]

Then given the fact that \( M(t) = N(t) - \int_{0}^{\tau} Y(u) \exp(\beta'Z) d\Lambda_0(u) \) is a mean zero martingale, Equation (6) is implied by the above two equations.

**References**