Received: 19 December 2018 | Revised: 18 May 2020 | Accepted: 1 June 2020

DOI: 10.1111/biom.13318

BIOMETRIC METHODOLOGY

Biometrics A JOURNAL OF THE INTERNATIONAL BIOMETRIC SOCIETY WILEY

Semiparametric regression calibration for general hazard models in survival analysis with covariate measurement error; surprising performance under linear hazard

Ching-Yun Wang[1] | **Xiao Song**[2]

[1] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington

[2] Department of Epidemiology and Biostatistics, College of Public Health, University of Georgia, Athens, Georgia

Correspondence
Ching-Yun Wang, Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, P.O. Box 19024, Seattle, WA 98109.
Email: cywang@fredhutch.org

Funding information
National Cancer Institute, Grant/Award Numbers: CA235122, CA239168, CA201207; National Heart, Lung, and Blood Institute, Grant/Award Number: HL130483; National Science Foundation, Grant/Award Number: 1916411

Abstract

Observational epidemiological studies often confront the problem of estimating exposure-disease relationships when the exposure is not measured exactly. Regression calibration (RC) is a common approach to correct for bias in regression analysis with covariate measurement error. In survival analysis with covariate measurement error, it is well known that the RC estimator may be biased when the hazard is an exponential function of the covariates. In the paper, we investigate the RC estimator with general hazard functions, including exponential and linear functions of the covariates. When the hazard is a linear function of the covariates, we show that a risk set regression calibration (RRC) is consistent and robust to a working model for the calibration function. Under exponential hazard models, there is a trade-off between bias and efficiency when comparing RC and RRC. However, one surprising finding is that the trade-off between bias and efficiency in measurement error research is not seen under linear hazard when the unobserved covariate is from a uniform or normal distribution. Under this situation, the RRC estimator is in general slightly better than the RC estimator in terms of both bias and efficiency. The methods are applied to the Nutritional Biomarkers Study of the Women's Health Initiative.

KEYWORDS
instrumental variable, measurement error, surrogate, survival analysis

1 | INTRODUCTION

Estimation of exposure-disease relationships in epidemiological studies may encounter the challenge of exposure measurement error. This is especially common when the exposure is quantitative and must be measured or estimated from characteristics of the individual and/or circumstances of exposure. Some of the most important examples of this problem arise in nutrient intake, physical activity, radiation, and other environmental exposures. It is widely recognized that errors or uncertainties in exposure variables can introduce bias into estimates of exposure-disease relationships.

Regression calibration (RC) is a statistical method for adjusting regression coefficient estimation for bias due to measurement error in exposure variables. The RC method for covariate measurement error is to replace an error-prone covariate by its conditional expectation given the observed covariates. In linear regression, RC is a consistent estimator for regression coefficients (Buonaccorsi, 2010, chapter 5). However, for logistic and Cox regression, it is known that it is not consistent (Carroll *et al.*, 2006,

chapter 4). There is further research on refinement of RC for logistic and Cox regression; see, for example, Wang *et al.* (2000).

An important covariate measurement error application is dietary intake as a risk factor for disease. In the past, self-report data are used as a tool for dietary intake. Recently, dietary biomarker studies have been proposed to understand the measurement biases associated with self-report data. For example, the Nutrient Biomarker Study (NBS) within the Dietary Modification Trial component of the Women's Health Initiative (WHI). In 2004-2005, the NBS recruited 544 subjects among 12 WHI clinical centers. Doubly labeled water was used for assessment of energy consumption and urinary nitrogen assessment of protein consumption (Neuhouser *et al.*, 2008). The statistical models proposed in Sugar *et al.* (2007) can accommodate a systematic error term that is allowed to depend on personal characteristics. These papers assumed that biomarker data may adhere to a classical measurement error model, whereas self-report data are linearly correlated with the underlying nutrient intake of interest. Another dietary biomarker study was the National Cancer Institute's Observing Protein and Energy Nutrition (OPEN) Study, which involved doubly labeled water and urinary nitrogen assessments, along with questionnaires and 24-hour recalls for 261 men and 223 women in Maryland. See Kipnis *et al.* (2001, 2003) for measurement error modeling for the OPEN study. In addition to dietary data, measurement error in biomarkers may cause estimation bias. For example, in HIV/AIDS research, CD4 lymphocyte count is an important biomarker for functionality of the immune system. However, CD4 count may contain measurement errors since it has no gold standard measurement and may contain biological fluctuation (Wu et al., 2008, 2010). Methodology for covariate measurement error with a flexible error model would improve effect estimation in many studies (Song and Wang, 2014).

Although an exponential hazard function has been popular (Qi *et al.*, 2005), there are situations when a linear hazard function may be a better fit (such as radiation effects). Therefore, we are motivated to develop methodology for measurement error for a general class of hazard functions. In this paper, we propose a semiparametric RC and risk set RC (RRC) estimators in survival analysis under general hazard functions. When the hazard function is linear, we show that the RRC estimator is consistent and robust to a working linear model for the unobserved exposure given the observed covariates at each risk set. We present a surprising finding that under a linear hazard function, the trade-off between bias and efficiency in measurement error research does not hold when comparing the RC and RRC estimators if the unobserved covariate is from a uniform or normal distribution. In Section 2, we describe the regression models in our problem. In Section 3, we review RC for Cox regression with measurement error. In Section 4, we investigate a semiparametric RC estimator when the hazard is a linear function of the covariates. The performance of the RC estimator is investigated in Section 5, and is compared with the RRC estimator. We apply the methods to the NBS data in Section 6 to study the association between protein intake and breast cancer. Some concluding remarks are given in Section 7. Technical proofs are given in the Web Appendix in the Supplementary Materials.

2 | STATISTICAL MODELS

In our problem of interest, we assume that the study cohort consists of n subjects. For $i = 1, \dots, n$, let X_i be the primary but unobserved exposure variable that may be associated with a disease outcome. For example, X_i may be dietary intake, radiation exposure, or physical activity in an epidemiological study. We assume that X_i is a scalar variable for notational simplicity. We assume that there is a surrogate measurement for X that follows the classical additive measurement error model

$$W_i = X_i + U_i, \text{with } \mathrm{E}(U_i|X_i) = 0, \tag{1}$$

for $i = 1, \dots, n$. Here we assume there is only one W_i, but the methods can be easily applied to the situation when replicates are available. For example, in diet and disease studies, biomarker-measured nutrient (such as doubly labeled water and urinary nitrogen assessments) may be considered as an unbiased surrogate that follows the additive measurement error model (1) given above. Let $\boldsymbol{Z}_i$ be a vector of covariates measured without an error, such as age, gender, and body mass index. Let T_i^0 be the survival time of the ith subject, and C_i be the censoring time. The response consists of observed variables $T_i \equiv \min(T_i^0, C_i)$ and $\delta_i \equiv I(T_i^0 \le C_i)$, where $I(\cdot)$ is the indicator function. Of interest is the relationship between survival time T_i^0 and covariates X_i and $\boldsymbol{Z}_i$, but T_i^0 is subject to censoring and thus is not fully observed. We assume that T_i^0 is independent of C_i given X_i and $\boldsymbol{Z}_i$, and the cumulative hazards function Λ of T_i^0 given $(X_i, \boldsymbol{Z}_i)$ follows the following general hazard model

$$\Lambda(dt|X_i, \boldsymbol{Z}_i) = \Lambda_0(dt) r(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i), \tag{2}$$

where $\boldsymbol{\beta}$ is a vector of parameters of interest, $\Lambda_0(\cdot)$ is an unspecified baseline cumulative hazard function, and $r(\boldsymbol{\beta}, X, \boldsymbol{Z})$ is the relative risk function. The hazard model (2) includes the Cox (1972) proportion model when $r(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i) = \exp(\beta_1 X_i + \boldsymbol{\beta}_2' \boldsymbol{Z}_i)$. A linear hazard model

such as $r(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i) = 1 + \beta_1 X_i + \boldsymbol{\beta}_2' \boldsymbol{Z}_i$ was investigated in Thomas (1981) and Prentice and Mason (1986). A less attractive property of the linear hazard model is that the hazard function may be negative at some parameter values and some ranges of the covariates. Similar to the linear hazard model given above, Wang *et al.* (2017) investigated a linear excess relative risk (ERR) model with $r(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i) = \exp(\boldsymbol{\beta}_2' \boldsymbol{Z}_i)\{1 + \beta_1 X_i \exp(\boldsymbol{\beta}_3' \boldsymbol{Z}_i)\}$. In the ERR model, β_1 is ERR per unit dose for the exposure at baseline ($Z = 0$), $\boldsymbol{\beta}_2$ is to model the background disease rate as a function of covariate $\boldsymbol{Z}$, and $\boldsymbol{\beta}_3$ is for the ERR effect modification by $\boldsymbol{Z}$.

We assume that there is an instrumental variable (IV) that is associated with the unobserved true exposure variable. Roughly speaking, a variable is an IV if it is correlated with the unobserved exposure, independent of the measurement error of the surrogate variable for the true exposure, and independent of the outcome variable given the covariates X_i and $\boldsymbol{Z}_i$. We assume that the IV, Q_i, follows the following general model:

$$Q_i = h(X_i, \boldsymbol{Z}_i) + V_i, \tag{3}$$

where $h(X, \boldsymbol{Z})$ can be any unknown function, or can be a known function but with unknown parameters, and V_i is a random error such that $\mathrm{E}(V_i | X_i, \boldsymbol{Z}_i) = 0$. For example, $Q_i = \gamma_0 + \gamma_1 X_i + \boldsymbol{\gamma}_2' \boldsymbol{Z}_i + \boldsymbol{\gamma}_3' X_i \boldsymbol{Z}_i + V_i$, where $\gamma_0, \gamma_1, \boldsymbol{\gamma}_2$, and $\boldsymbol{\gamma}_3$ are unknown coefficients. In (3), V_i has mean 0 and is independent of the variables in (3). The measurement error $Q - X$ in (3) is subject-specific as it is related to an individual's characteristics $\boldsymbol{Z}_i$. As discussed in the introduction, the self-report dietary data may be associated with a subjective or systematic bias and hence it may not follow well the additive measurement error model (1). Instead, the flexible model (3) will likely hold for self-report nutrient data. Here we assume there is only one Q_i for each subject, but the methods to be developed later can be applied to the situation when replicates are available.

3 | RC FOR COX REGRESSION

In this section, we briefly review and discuss why RC in Cox regression with measurement error is not consistent. For $i = 1, \ldots, n$, let T_i^0 be the survival time of the ith subject, and C_i be the censoring time. The response consists of observed variables $T_i \equiv \min(T_i^0, C_i)$ and $\delta_i \equiv I(T_i^0 \le C_i)$, where $I(\cdot)$ is the indicator function. Of interest is the relationship between survival time T^0 and covariates X_i and $\boldsymbol{Z}_i$, but T_i^0 is subject to censoring and thus is not fully observed. We assume that T_i^0 is independent of C_i given X_i and $\boldsymbol{Z}_i$, and the hazard function λ of T_i^0 given $(X_i, \boldsymbol{Z}_i)$ follows the Cox proportional hazards model

$$\lambda(t; X_i, \boldsymbol{Z}_i) = \lambda_0(t) \exp(\beta_1 X_i + \boldsymbol{\beta}_2' \boldsymbol{Z}_i),$$

where $\lambda_0(t)$ is the unspecified baseline hazard function. The hazard function given above cannot be applied directly to the estimating procedure as X is not available. Because the surrogate variable W for X is available, one approach to address the measurement error issue is to derive the *induced* hazard function given the observed data. As in Prentice (1982), the induced hazards function could be approximated by the following:

$$\begin{aligned} &\mathrm{E}(e^{\beta_1 X_i + \boldsymbol{\beta}_2' \boldsymbol{Z}_i} | \boldsymbol{Z}_i, Q_i, T_i \ge t) \\ &\quad \approx e^{0.5\beta_1^2 \mathrm{var}(X_i | \boldsymbol{Z}_i, Q_i, T_i \ge t)} e^{\beta_1 \mathrm{E}(X_i | \boldsymbol{Z}_i, Q_i, T_i \ge t) + \boldsymbol{\beta}_2' \boldsymbol{Z}_i}, \end{aligned} \tag{4}$$

where the approximation given above was based on a Taylor expansion. Hence, the RC estimator would have limited bias if (a) the measurement error is from a normal distribution; (b) $\beta_1^2 \mathrm{var}(X_i | W_i, Q_i, \boldsymbol{Z}_i)$ is small; and (c) the disease is rare. In Cox regression with covariate measurement error under the setup with replicates in W, Xie *et al.* (2001) showed that the RRC estimator has smaller biases than the RC estimator in general. The RRC estimator would replace X_i by $\mathrm{E}(X_i | \boldsymbol{Z}_i, Q_i, T_i \ge t)$ in solving the usual estimating equation (from partial likelihood in Cox regression). However, they also showed that the RRC estimator still could have a bias problem under some situations as it is not a consistent estimator.

In measurement error research the trade-off between bias and efficiency is often referred to the comparison between the naive estimator and a consistent estimator, but it is also applicable to a comparison between the RC and RRC estimators in nonlinear regression. In Cox regression, the trade-off between bias and efficiency was also noted between the naive estimator, RC and RRC estimators (Xie *et al.*, 2001).

A parametric RC estimator may assume that the joint distribution of X, W, Q, and $\boldsymbol{Z}$ is multivariate-normal (Carroll *et al.*, 2006, chapter 4) if these variables are continuous. If $\boldsymbol{Z}$ is discrete, then the RC can be implemented by assuming that the joint conditional distribution of X, W, Q given $\boldsymbol{Z}$ is multivariate-normal. With this model assumption, $\mathrm{E}(X_i | W_i, Q_i, \boldsymbol{Z}_i)$, or $\mathrm{E}(X_i | Q_i, \boldsymbol{Z}_i)$, could serve as a replacement for X_i. However, the parametric RC estimator is generally somewhat computationally complicated. Hence, we will consider a semiparametric RC estimator. The semiparametric RC estimator is to replace X by modeling regression of W given $(Q, \boldsymbol{Z})$, and then $\mathrm{E}(W | Q, \boldsymbol{Z})$ is estimated by the predicted value of W given $(Q, \boldsymbol{Z})$. For

example, by checking the observed data, we could model the relation between W and $(Q, \boldsymbol{Z})$ by $\mathrm{E}_*(W|Q,\boldsymbol{Z}) = \alpha_0 + \alpha_1 Q_i + \boldsymbol{\alpha}_2' \boldsymbol{Z}_i$ with parameters $\alpha_0, \alpha_1, \boldsymbol{\alpha}_2$ that can be estimated by least square estimation. Here we use the notation E_* to indicate that the conditional expectation is based on the working model rather than the true conditional expectation of W given $(Q, \boldsymbol{Z})$. Here $\mathrm{E}_*(W|Q,\boldsymbol{Z})$ is the same as $\mathrm{E}_*(X|Q,\boldsymbol{Z})$ as W is an unbiased surrogate for X.

A semiparametric RRC estimator can be implemented by calculating $\mathrm{E}(W|Q,\boldsymbol{Z}, T \geq t)$ for $\mathrm{E}(X|Q,\boldsymbol{Z}, T \geq t)$. In each risk set, we could consider a working regression model for W given $(Q, \boldsymbol{Z})$. For example, within each risk set, W may be modeled as linear, such as $\mathrm{E}_*(W|Q,\boldsymbol{Z}, T \geq t) = \alpha_{0t} + \alpha_{1t} Q + \boldsymbol{\alpha}_{2t}' \boldsymbol{Z}$ if the association is appropriate. From the induced hazard function (4), the RC and RRC estimators may be different if (a) the measurement error is not from a normal distribution; (b) $\beta_1^2 \mathrm{var}(X_i|W_i, Q_i, \boldsymbol{Z}_i)$ is large; and (c) the disease is not rare. Among the three factors, the magnitude of $|\beta_1|$ is usually the most important. If the standard deviation of the covariates is about 1 and $|\beta_1|$ is larger than ln(2), then the RRC estimator usually has smaller bias than that from the RC estimator, but RC could be slightly better if the event rate is low (say less than 10%).

4 | RC FOR LINEAR HAZARD REGRESSION

In this section, we investigate an RC estimator when the hazard is a linear function of the covariates such that $r(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i) = 1 + \beta_1 X_i + \boldsymbol{\beta}_2' \boldsymbol{Z}_i$. The model is slightly different from a linear ERR model with $r(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i) = \exp(\boldsymbol{\beta}_3' \boldsymbol{Z}_i)\{1 + \beta_1 X_i \exp(\boldsymbol{\beta}_2' \boldsymbol{Z}_i)\}$, but the methodology development will be similar. The partial likelihood score estimating equation for the general hazard model in the absence of measurement error can be written as

$$n^{-1/2} \sum_{i=1}^{n} \int_0^{\tau} \left\{ \frac{r^{(1)}(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i)}{r(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i)} - \frac{\sum_{j=1}^{n} Y_j(t) r^{(1)}(\boldsymbol{\beta}, X_j, \boldsymbol{Z}_j)}{\sum_{j=1}^{n} Y_j(t) r(\boldsymbol{\beta}, X_j, \boldsymbol{Z}_j)} \right\} \times dN_i(t) = \mathbf{0},$$

where $r^{(1)}(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i) = (\partial/\partial\boldsymbol{\beta}) r(\boldsymbol{\beta}, X_i, \boldsymbol{Z}_i)$ is the derivative of the relative risk function with respect to $\boldsymbol{\beta}$, and τ is the time limit. As discussed in Thomas (1981) and Prentice and Mason (1986), the partial likelihood score given above may encounter finite sample challenges as the relative risk function is involved in the denominator of the estimating equation. To avoid this issue, the following estimating equation (when there is no measurement error) can be shown to be unbiased.

$$\boldsymbol{\Psi}_n(\boldsymbol{\beta}, X, \boldsymbol{Z}) \equiv n^{-1/2} \sum_{i=1}^{n} \int_0^{\tau} \left\{ \begin{pmatrix} X_i \\ \boldsymbol{Z}_i \end{pmatrix} - \frac{\sum_{j=1}^{n} Y_j(u)(X_j, \boldsymbol{Z}_j')' r(\boldsymbol{\beta}, X_j, \boldsymbol{Z}_j)}{\sum_{j=1}^{n} Y_j(u) r(\boldsymbol{\beta}, X_j, \boldsymbol{Z}_j)} \right\} \times dN_i(u) = \mathbf{0}. \quad (5)$$

Estimating Equation (5) given above can be written as a martingale representation and hence it can be called a martingale-based estimating equation (MEE, Wang *et al.*, 2017). When X is measured with an error, we may calculate the expected value of the hazard function given the observed data, namely the *induced hazard function*. It can be seen that the induced hazard function can be expressed as

$$\lambda(t|\boldsymbol{Z}_i, Q_i) = \lambda_0(t)\{1 + \beta_1 \mathrm{E}(X_i|\boldsymbol{Z}_i, Q_i, T_i \geq t) + \boldsymbol{\beta}_2' \boldsymbol{Z}_i\}.$$

From the equation given above, a consistent estimator for the model with measurement error can be obtained by replacing X_i with $\mathrm{E}(X_i|Q_i, \boldsymbol{Z}_i, T_i \geq t)$ in the estimating equation. That is, the RRC estimator is consistent under a linear hazard function. When the at risk indicator $T_i \geq t$ in the calculation of $\mathrm{E}(X_i|\boldsymbol{Z}_i, Q_i, T_i \geq t)$ is ignored, this is the RC estimator when X_i is replaced by $\mathrm{E}(X_i|\boldsymbol{Z}_i, Q_i)$. The RC can be implemented by calculating $\mathrm{E}(W_i|\boldsymbol{Z}_i, Q_i)$ as an approximation for $\mathrm{E}(X|\boldsymbol{Z}_i, Q_i)$.

We now investigate potential differences between $\mathrm{E}(X_i|Q_i, \boldsymbol{Z}_i, T_i \geq t)$ and $\mathrm{E}(X_i|Q_i, \boldsymbol{Z}_i)$ in order to have insight regarding the differences between the RC and RRC estimators. By some calculations given in Equation (3) of the Web Appendix, under a special case when X is the only covariate and if X given Q is normally distributed (if X values satisfy $1 + \beta X > 0$), it can be seen that

$$\mathrm{E}(X|Q, T \geq t) = \mathrm{E}(X|Q) - \Lambda_0(t)\beta_1 \mathrm{var}(X|Q),$$

where $\Lambda_0(\cdot)$ denotes the baseline cumulative hazards function, and $\mathrm{var}(X|Q)$ is the conditional variance of X given Q. Hence, the RC and RRC estimators are different under this special case.

As mentioned in the previous section, by checking the observed data, we could model the relation between W and $(Q, \boldsymbol{Z})$ by $\mathrm{E}_*(W|Q,\boldsymbol{Z}) = \alpha_0 + \alpha_1 Q + \boldsymbol{\alpha}_2' \boldsymbol{Z}$ with parameters $\alpha_0, \alpha_1, \boldsymbol{\alpha}_2$. This model is a working model, and it may not hold under the model in (3) such that $Q_i = h(X_i, \boldsymbol{Z}_i) + V_i$. For example, if Q given X is linear in X, then X given Q may not be linear in Q, and the association may be

more complicated within each risk set. From our simulation study, interestingly the RRC estimator is not sensitive to the working model assumption. For example, the RRC estimator has limited biases (when comparing with the standard errors) even when Q given X is quadratic while the working model of W given Q is assumed to be linear, and the bias decreases to 0 when the sample size increases. This motivates our investigation on the robust property of the RRC estimator. At each risk set $T_i \geq u$, we assume a working model $\mathrm{E}_*(W_i|Q_i, \boldsymbol{Z}_i, T_i \geq u) = \alpha_{0u} + \alpha_{1u}Q_i + \boldsymbol{\alpha}'_{2u}\boldsymbol{Z}_i$ such that

$$\sum_{i=1}^{n} Y_i(u)\begin{pmatrix}1\\ Q_i\\ \boldsymbol{Z}_i\end{pmatrix}\{W_i - \alpha_{0u} - \alpha_{1u}Q_i - \boldsymbol{\alpha}'_{2u}\boldsymbol{Z}_i\} = 0, \quad (6)$$

where $\boldsymbol{\alpha}_u = (\alpha_{0u}, \alpha_{1u}, \boldsymbol{\alpha}'_{2u})'$ is a vector of parameters. The notation E_* is used to indicate that the conditional expectation is based on the working model rather than the true conditional expectation of W given $(Q, \boldsymbol{Z})$. At each risk set, the RRC based on the working model replaces X_i with $\widehat{X}^*_i(u) \equiv \widehat{\alpha}_{0u} + \widehat{\alpha}_{1u}Q_i + \widehat{\boldsymbol{\alpha}}'_{2u}\boldsymbol{Z}_i$ where $\widehat{\boldsymbol{\alpha}}$ solves (6). The estimating equation for the RRC estimator can be expressed as $\boldsymbol{\Psi}_n(\boldsymbol{\beta}, \widehat{X}^*(u), \boldsymbol{Z}) = 0$, where $\boldsymbol{\Psi}_n(\boldsymbol{\beta}, X, \boldsymbol{Z})$ is given in (5). Let the proposed semiparametric RRC estimator be denoted by $\widehat{\boldsymbol{\beta}}_{rrc}$.

Proposition 1. *Assume that the relative risk function is linear with $r(\boldsymbol{\beta}, X, \boldsymbol{Z}) = 1 + \beta_1 X + \boldsymbol{\beta}'_2\boldsymbol{Z}$. The surrogate variable W satisfies the classical additive measurement error model (1), and the IV Q satisfies a general model (3). At each risk set, we assume a working model (6) to replace X. Then $\widehat{\boldsymbol{\beta}}_{rrc}$ converges to $\boldsymbol{\beta}$ in probability, and $n^{1/2}(\widehat{\boldsymbol{\beta}}_{rrc} - \boldsymbol{\beta})$ is asymptotically normal with mean 0 and variance given in Section 3 (the Web Appendix) of the Supplementary Materials.*

The working model $\mathrm{E}_*(W|Q, \boldsymbol{Z}, T \geq u) = \alpha_{0u} + \alpha_{1u}Q + \boldsymbol{\alpha}'_{2u}\boldsymbol{Z}$ can be modified to another more suitable regression model in the application, and the robustness of the RRC estimator still holds. Proposition 1 can be extended to the ERR model described earlier, and the robustness of the RRC estimator to the working model still holds. The proof of Proposition 1 is given in Section 3 of the Supplementary Materials.

Under linear hazard regression, the RC and RRC estimators are equally well overall. They may be close numerically under some situations, but one may be better than the other under other situations. If the covariate distribution is symmetric (such as uniform or normal), then in general the RRC estimator is slightly better than the RC estimator in terms of both bias and efficiency (Tables 2 and 3, in where $n = 400$ and 800), but they could be very close (Table 2). If the covariate distribution is very skewed (Table 4, $n = 900$ and 1300) and the event rate is 10%, then the RC estimator has better finite sample performance than the RRC estimator. However, if the event rate is 60%, there is trade-off between bias and efficiency; the RRC estimator has smaller bias and better coverage probabilities than the RC estimator.

5 | SIMULATION STUDY

We conducted a simulation study to evaluate the performance of the semiparametric RC estimator. The naive estimator is to use W to replace X. We considered the RC estimator that replaces X by a working model that $\mathrm{E}_*(W|Q, \boldsymbol{Z}) = \alpha_0 + \alpha_1 Q + \boldsymbol{\alpha}'_2\boldsymbol{Z}$. The RRC estimator replaces X by $E(X|Q, \boldsymbol{Z}, T \geq u)$ in each risk set based on a working model $\mathrm{E}_*(W|Q, \boldsymbol{Z}, T \geq u) = \alpha_{0u} + \alpha_{1u}Q + \boldsymbol{\alpha}'_{2u}\boldsymbol{Z}$ in each risk set. In Table 1, the covariates X are from a normal distribution with $\mu_x = 1$, $\sigma_x = 1$. The surrogates $W_i, i = 1, \ldots, n$ are generated by $W_i = X_i + U_i$, where U_i is normal with mean 0 and standard deviation $\sigma_u = 0.5$. The IVs are generated based on $Q_i = \gamma_0 + \gamma_1 X_i + V_i$, where V_i is normal with mean 0 and standard deviation $\sigma_v = 0.5$, and $\gamma_0 = 0.2$, $\gamma_1 = 0.7$. The failure times are generated by the hazard function $\lambda(t; X_i) = \exp(\beta X_i)$, where $\beta = \ln(1.5)$ and $\ln(2)$, respectively. The censoring time is a fixed time such that the event rate is 5% and 50%, respectively. The sample size of the whole cohort in the simulation is $n = 400$ and $n = 800$, respectively. In the tables, "bias" is the average of $\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}$ from 500 replicates, "SD" denotes the sample standard deviation of the estimators, "ASE" denotes the average of the estimated standard errors of the estimators. The 95% Wald-type confidence interval coverage probabilities are also included. All the parameters are given in the tables. The standard errors (SEs) of the RC estimates are based on a sandwich variance estimator where the vector of the estimating equations is obtained by stacking the estimating equations for $\boldsymbol{\beta}$ and the estimating equations for regression W given Q discussed in Section 4. The SEs of the semiparametric RRC estimates are based on the sandwich variance estimator given in the Web Appendix. From the simulation result of Table 1, it is seen that the naive estimator has large biases. The RC estimator performs reasonably well in terms of bias correction when $\beta = \ln(1.5)$, but the biases are larger when $\beta = \ln(2)$. The RRC estimator has similar performance as RC when $\beta = \ln(1.5)$. Under the exponential relative risk function, the RRC estimator is able to reduce the biases of the RC estimator when the relative risk parameter increases to $\beta = \ln(2)$. From this table, the main finding is that in Cox regression for a small relative risk parameter such as $\ln(1.5)$ the RC

TABLE 1 Simulation under exponential hazard function with normal X

		$n = 400$			$n = 800$		
		Naive	**RC**	**RRC**	**Naive**	**RC**	**RRC**
$\beta = \ln(1.5)$							
event rate = 0.05							
β	Bias	−0.193	0.023	0.025	−0.203	0.008	0.008
	SD	0.159	0.290	0.291	0.115	0.182	0.183
	ASE	0.155	0.270	0.271	0.110	0.193	0.193
	CP	0.742	0.922	0.920	0.550	0.960	0.962
event rate = 0.50							
β	Bias	−0.208	−0.004	0.005	−0.208	−0.004	0.003
	SD	0.051	0.093	0.098	0.039	0.064	0.067
	ASE	0.051	0.091	0.095	0.036	0.065	0.067
	CP	0.022	0.950	0.946	0.002	0.944	0.936
		Naive	RC	RRC	Naive	RC	RRC
$\beta = \ln(2)$							
event rate = 0.05							
β	Bias	−0.340	0.025	0.030	−0.351	0.001	0.004
	SD	0.161	0.286	0.289	0.113	0.189	0.190
	ASE	0.156	0.273	0.276	0.111	0.194	0.195
	CP	0.432	0.950	0.948	0.108	0.950	0.948
event rate = 0.50							
β	Bias	−0.370	−0.034	−0.005	−0.372	−0.032	−0.004
	SD	0.053	0.098	0.110	0.040	0.069	0.076
	ASE	0.052	0.098	0.107	0.037	0.069	0.076
	CP	0.000	0.910	0.936	0.000	0.908	0.938

Note. The naive estimator replaced the unobserved X by W, the RC estimator replaced X by $E(W|Q)$. The RRC estimator replaced X by $E(X|Q, T \geq t)$ in each risk set. Parameters are $\mu_x = 1$, $\sigma_x = 1$, $\sigma_u = 1$. In addition, $Q_i = \gamma_0 + \gamma_1 X_i + V_i$, where $\gamma_0 = 0.2$, $\gamma_1 = 0.7$, and $\sigma_v = 0.5$. Results are from 500 replicates.

estimator is as good as the RRC estimator, and is slightly better when the event rate is 5%. When $\beta = \ln(2)$ with 50% event rate, the RRC estimator has smaller biases but with the cost of being less efficient.

In Table 2, we conducted analysis under a linear hazard model. The covariates variable X is from a uniform distribution, with $\mu_x = 1$, $\sigma_x = 1$. The surrogates W_i and IV Q_i are generated similarly to those in Table 1 with the parameters given in the table. The failure times are generated by the hazard function $\lambda(t; X_i) = 1 + \beta X_i$, where $\beta = 0.2$ and 0.4, respectively. The censoring time is a fixed time such that the event rate is 10% and 50%, respectively. From the development of the methods, the RRC estimator is a consistent estimator under linear hazard functions. The results in Table 2 indicates that the RRC estimator in most cases is very close to the RC estimator even though it is slightly better in terms of bias and efficiency. One surprising finding is that the trade-off between bias and efficiency in measurement error research is not seen under the linear hazard model in this table. Although not reported in the table, we did similar simulations but with normal X, and the result is similar.

In Table 3, we investigate the situation similar to Table 2 but the relation between Q and X is no longer linear, with $Q_i = \gamma_0 + \gamma_1 X_i^2 + V_i$, where $\gamma_0 = 0.1$, $\gamma_1 = 0.95$, and $\sigma_v = 0.25$. Also, in this table the measurement error U is from a mixture of two normal variables; one with mean 1.2 and variance 1.44, the other with mean −0.6 and variance 0.36, and the mixture percentages are (2/3, 1/3). The measurement error in this table has an asymmetric distribution. The RC estimator is based on a linear working model such that X is replaced with $E_*(W|Q) = \alpha_0 + \alpha_1 Q$. The RRC estimator replaces X with linear $E(X|Q, T \geq t)$ in each risk set. The relationship between Q and X is no longer linear, and the working model for W given Q is a misspecification of the true association between W and Q. Hence, this table demonstrates the situation when the working model is very different from the true model. From this table, it is seen that the results of Table 3 are similar to those of Table 2. The RC and RRC estimators are both robust to the linear working model of W given Q. The RRC estimator has small advantages over the RC estimator in terms of bias and efficiency, and the advantage can be seen better with a larger β_1. In order to understand the surprising

TABLE 2 Simulation under linear hazard function with uniform X

		$n = 400$			$n = 800$		
		Naive	**RC**	**RRC**	**Naive**	**RC**	**RRC**
$\beta = 0.2$							
event rate = 0.10							
β	Bias	−0.085	0.087	0.087	−0.092	0.037	0.037
	SD	0.139	0.365	0.364	0.098	0.206	0.205
	ASE	0.139	0.318	0.318	0.097	0.198	0.199
	CP	0.810	0.950	0.950	0.748	0.948	0.948
event rate = 0.50							
β	Bias	−0.104	0.025	0.024	−0.105	0.008	0.008
	SD	0.057	0.128	0.128	0.042	0.089	0.089
	ASE	0.060	0.122	0.123	0.042	0.083	0.083
	CP	0.550	0.952	0.956	0.312	0.930	0.932
		Naive	RC	RRC	Naive	RC	RRC
$\beta = 0.4$							
event rate = 0.10							
β	Bias	−0.198	0.172	0.167	−0.216	0.049	0.048
	SD	0.161	0.965	0.902	0.111	0.285	0.284
	ASE	0.159	0.600	0.573	0.109	0.268	0.268
	CP	0.642	0.968	0.968	0.454	0.938	0.940
event rate = 0.50							
β	Bias	−0.226	0.037	0.034	−0.227	0.013	0.011
	SD	0.064	0.174	0.173	0.047	0.123	0.123
	ASE	0.067	0.165	0.165	0.047	0.111	0.111
	CP	0.108	0.956	0.958	0.008	0.934	0.934

Note. The naive estimator replaces the unobserved X with W, the RC estimator replaces X with $E(W|Q)$. The RRC estimator replaces X with $E(X|Q, T \geq t)$ in each risk set. Parameters are $\mu_x = 1, \sigma_x = 1, \sigma_u = 1$. In addition, $Q_i = \gamma_0 + \gamma_1 X_i + V_i$, where $\gamma_0 = 0.1, \gamma_1 = 0.9$, and $\sigma_v = 0.5$. Results are from 500 replicates.

nontrade-off between bias and efficiency when comparing the RC and RRC estimators under linear hazard functions, we calculated the biases and SDs of the RC and RRC with a series of sample sizes. Web Figure 1 of the Supplementary Materials shows the biases and SDs for both exponential and linear hazard functions, respectively. Under the exponential hazard model, it is clear that the RRC estimates have smaller biases than the RC estimates, but the RRC has the larger variations. Under linear hazard, the RRC bias decreases to 0 when the sample size increases, but the RC bias remains about the same with increasing sample sizes. This is basically because under linear hazard models, the RRC estimator is consistent, which is not the case for the RC estimator. Nevertheless, finite sample performance of the RC estimator could be almost as good as the RRC estimator under many practical situations.

We further investigate the situation when the error-prone covariate distribution is skewed. In Table 4, we generated the data similarly to Table 2, but the covariates $X_i, i = 1, \ldots, n$, were from a shifted scaled chi-square distribution with mean $\mu_x = 1$ and variance $\sigma_x^2 = 1$, and the measurement error U_i is normal or shifted scaled chi-square, respectively, with mean 0 and standard deviation $\sigma_u = 1$. The sample size was $n = 900$ and 1300, respectively. Some additional parameters are given in the table. It is seen that the naive estimator has large biases. If the event rate is 10%, then the RC estimator has better finite sample performance than the RRC estimator. However, if the event rate is 60%, there is trade-off between bias and efficiency; the RRC estimator has smaller bias and better coverage probabilities than the RC estimator.

Another interesting question regarding the RRC estimator is whether it is possible to estimate $\mathrm{E}(W|Q, \boldsymbol{Z}, T \geq u)$ parametrically. In general, it is not easy to estimate $\mathrm{E}(W|Q, \boldsymbol{Z}, T \geq u)$ parametrically. When $\mathrm{E}(W|Q, \boldsymbol{Z}, T \geq u)$ can be estimated based on the correct covariate distribution, a parametric RRC (PRRC) may be more efficient than the proposed semiparametric RRC estimator. In Section 3 of the Supplementary Materials, we investigate a PRRC estimator when $(X, W, Q, \boldsymbol{Z})$ is multivariate normal. Simulation results show that if (X, W, Q, Z) is multivariate normal, then the PRRC estimator is more efficient than the semiparametric RRC estimator. However, if (X, W, Q, Z) is not from a multivariate normal

TABLE 3 Simulation under linear hazard function with uniform X, but E($Q|X$) is a quadratic function, mixture-normal measurement error

		$n = 400$			$n = 800$		
		Naive	**RC**	**RRC**	**Naive**	**RC**	**RRC**
$\beta = 0.2$							
event rate = 0.10							
β	Bias	−0.115	0.066	0.063	−0.117	0.036	0.034
	SD	0.134	0.403	0.391	0.084	0.237	0.232
	ASE	0.120	0.332	1.172	0.084	0.208	0.206
	CP	0.708	0.912	0.910	0.634	0.922	0.922
event rate = 0.50							
β	Bias	−0.121	0.017	0.012	−0.125	0.009	0.004
	SD	0.055	0.129	0.124	0.036	0.087	0.083
	ASE	0.053	0.127	0.124	0.037	0.088	0.085
	CP	0.378	0.932	0.932	0.094	0.960	0.954
		Naive	RC	RRC	Naive	RC	RRC
$\beta = 0.4$							
event rate = 0.10							
β	Bias	−0.250	0.221	0.177	−0.255	0.072	0.065
	SD	0.151	1.755	1.137	0.093	0.424	0.398
	ASE	0.134	1.271	0.780	0.093	0.316	0.308
	CP	0.454	0.900	0.900	0.258	0.920	0.920
event rate = 0.50							
β	Bias	−0.260	0.040	0.021	−0.264	0.027	0.009
	SD	0.060	0.190	0.175	0.039	0.126	0.116
	ASE	0.058	0.186	0.174	0.040	0.127	0.119
	CP	0.030	0.942	0.932	0.000	0.960	0.950

Note. The naive estimator replaces the unobserved X with W, the RC estimator replaces X with $E(W|Q) = \alpha_0 + \alpha_1 Q$. The RRC estimator replaces X with $E(X|Q, T \geq t)$ in each risk set. Parameters are $\mu_x = 1$, $\sigma_x = 1$, $\sigma_u = 1.2$. In addition, $Q_i = \gamma_0 + \gamma_1 X_i^2 + V_i$, where $\gamma_0 = 0.1$, $\gamma_1 = 0.95$, and $\sigma_v = 0.25$. Results are from 500 replicates.

distribution, then the PRRC estimator based on the multivariate normality assumption may have moderate biases. There is trade-off between bias and efficiency when comparing the proposed semiparametric RRC and PRRC estimators (Web Table 2).

6 | NBS DATA ANALYSIS

In this section, we demonstrate the RC and RRC estimators via the NBS of the WHI, which was briefly described in the introduction. We are interested in the association between protein intake, obesity, and breast cancer incidence. A subject is defined as obese if her body mass index (weight in Kg divided by square of height in meter) is higher than 30 (by World Health Organization). There are some motivations for the data analysis. For example, Fontana *et al.* (2006) showed that low protein intake may reduce IGF-1, independent of body weight. High levels of IGF-1 have been linked to breast cancer, prostate cancer, and certain types of colon cancer. The analysis in this section will investigate an association between percent energy from protein intake and breast cancer risk, adjusted for obesity. In the NBS, a subject's protein intake can be obtained from urinary nitrogen (UN) or a food frequency questionnaire (FFQ). As discussed in the introduction, FFQ may be associated with subjective or systematic bias and hence it may not follow well an additive measurement error model. In comparison, urinary nitrogen protein is less likely to be associated with a systematic bias. The true but unobserved urinary nitrogen protein of an individual is the average over a designated period of time. Urinary nitrogen protein from one sample may be associated with random variations from day-to-day biological fluctuations or random errors from the measurement process. Hence, we assume that the observed urinary nitrogen follows an additive measurement error model. Here percent energy from protein via FFQ is considered as an IV that is linearly associated with the underlying long-term average. FFQ protein is a reasonable IV as (a) it is correlated with the true

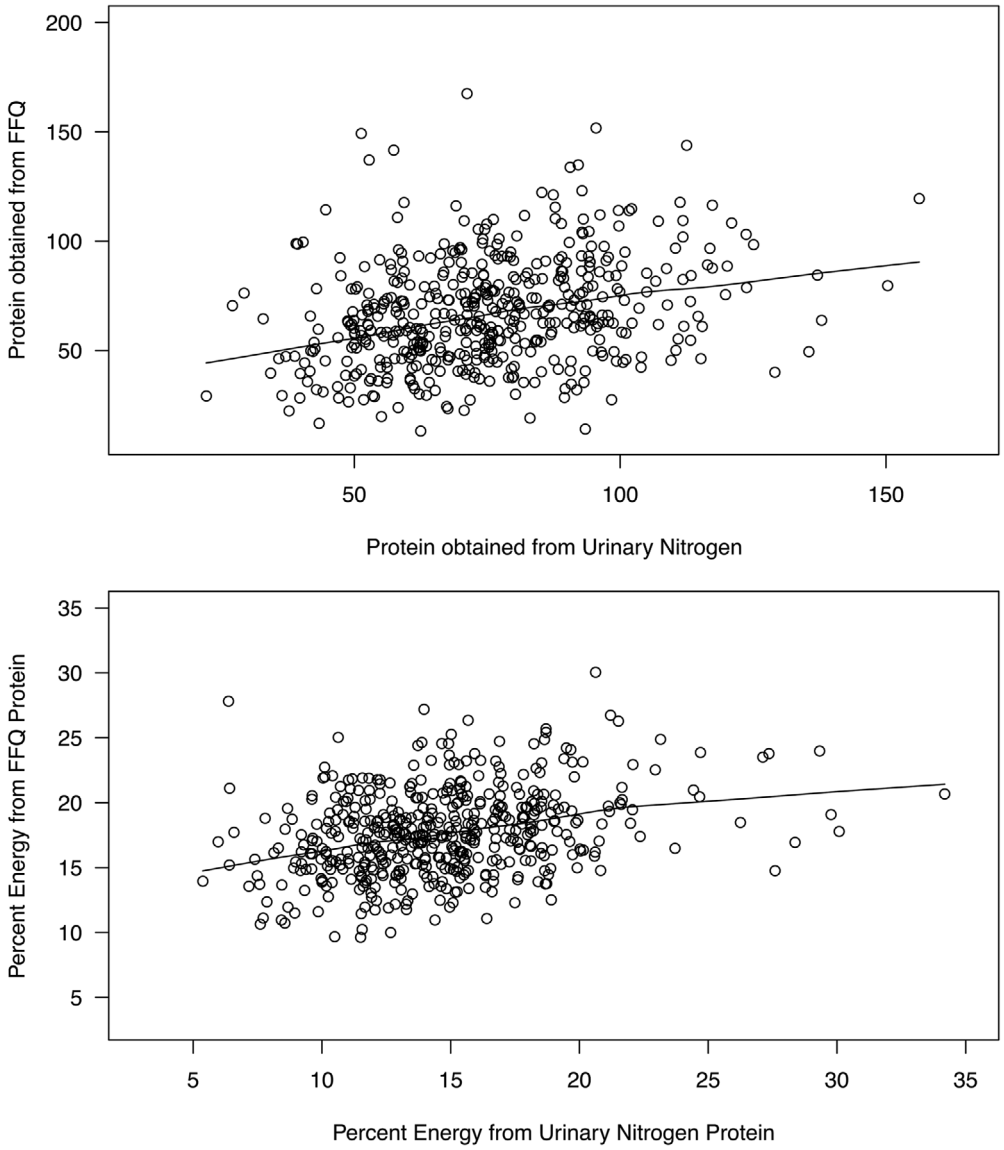

FIGURE 1 Scatterplots of protein from FFQ versus protein from urinary nitrogen, and % energy from protein via urinary nitrogen versus % energy from protein via FFQ, respectively. The lines are obtained from fitting lowess smoothers.

underlying protein intake; (b) it is not likely to be associated with the measurement error of urinary nitrogen protein; and (c) its effect on cancer risk is primarily due to the true underlying protein intake. In this application, 45 subjects of the NBS are not included as there were missing data in calculating their percent energy from protein either from FFQ or from UN. As a result, the analysis consists of 499 subjects who have both percent energy from protein via urinary nitrogen biomarker and FFQ. The WHI subjects were recruited during approximately 1994-1998, and the data used in the analysis were followed up until June 2005. In the NBS data, 18 individuals of the NBS are diagnosed with breast cancer by June 2005.

We now examine the surrogate and IV measurements in the analysis. The upper portion of Figure 1 is the scatter plot of protein intake obtained by UN versus that by FFQ, and a fitted lowess smoother. The protein intake data via UN and FFQ have moderate association with correlation coefficient of 0.30. The lower portion of Figure 1 is the scatter plot of percent energy from protein via UN versus that via FFQ. The correlation coefficient between percent energy from protein via UN versus that via FFQ is 0.33. From the figure, the association between protein via UN and protein via FFQ is reasonably linear. By comparing these two plots and the correlation coefficients, percent energy from protein data (UN and FFQ) are used in the analysis.

The development of breast cancer may be due to many environmental and genetic risk factors. But, the focus of our analysis is on the effect of protein intake and BMI in association with breast cancer risk. Hence, in our data analysis, percent energy from protein (unobserved

TABLE 4 Simulation under linear hazard function with shifted scaled chi-square X with mean 1 and variance 1

Measurement error U is normal with mean 0							
		$n = 900$			**$n = 1300$**		
		Naive	**RC**	**RRC**	**Naive**	**RC**	**RRC**
event rate = 0.10							
β	Bias	−0.222	0.111	0.145	−0.232	0.054	0.074
	SD	0.127	0.475	0.559	0.098	0.303	0.326
	ASE	0.123	0.426	0.486	0.101	0.306	0.330
	CP	0.498	0.916	0.918	0.358	0.918	0.922
event rate = 0.60							
β	Bias	−0.256	−0.044	0.029	−0.258	−0.049	0.019
	SD	0.042	0.115	0.165	0.038	0.099	0.138
	ASE	0.048	0.120	0.166	0.039	0.099	0.134
	CP	0.002	0.890	0.938	0.000	0.892	0.952
Measurement error U is shifted scaled chi-square with mean 0							
		$n = 900$			**$n = 1300$**		
		Naive	**RC**	**RRC**	**Naive**	**RC**	**RRC**
event rate = 0.10							
β	Bias	−0.216	0.100	0.139	−0.226	0.028	0.045
	SD	0.124	0.555	0.689	0.103	0.289	0.314
	ASE	0.127	0.442	0.527	0.102	0.294	0.316
	CP	0.508	0.918	0.924	0.396	0.940	0.942
event rate = 0.60							
β	Bias	−0.251	−0.038	0.037	−0.250	−0.050	0.017
	SD	0.049	0.129	0.183	0.043	0.099	0.136
	ASE	0.048	0.122	0.169	0.040	0.098	0.133
	CP	0.008	0.880	0.938	0.002	0.878	0.950

Note. The true β is 0.4. The naive estimator replaces the unobserved X with W, the RC estimator replaces X with $E(W|Q) = \alpha_0 + \alpha_1 Q$. The RRC estimator replaces X with $E(X|Q, T \geq t)$ in each risk set. Parameters are $\mu_x = 1$, $\sigma_x = 1$, and $\sigma_u = 1$. In addition, $Q_i = \gamma_0 + \gamma_1 X_i^2 + V_i$, where $\gamma_0 = 0.1$, $\gamma_1 = 0.9$, and $\sigma_v = 0.8$. Results are from 500 replicates.

long term average) and obesity are the covariates of interest. The data analysis in this section is primarily for demonstration of our new methods. We do not intend to interpret our findings as WHI results in dietary or obesity research. We analyzed the data based on Cox regression with $r(\boldsymbol{\beta}, X, \boldsymbol{Z}) = \exp(\beta_1 X + \beta_2 Z)$, linear hazard with $r(\boldsymbol{\beta}, X, \boldsymbol{Z}) = 1 + \beta_1 X + \beta_2 Z$, and EER $r(\boldsymbol{\beta}, X, \boldsymbol{Z}) = \exp(\beta_2 Z)\{1 + \beta_1 X \exp(\beta_3 Z)\}$. However, in the ERR analysis we assume $\beta_3 = 0$ as including the effect modification parameter in the model would encounter divergence. The issue of divergence when adding β_3 in the ERR analysis is due to the low event rate (18 cases), which would be more numerically challenging when the hazard function is linear, which was seen from our simulation study in the last section. The estimates from the naive, RC, and RRC estimators for each of the three hazard models are given in Table 5. The SEs of the RC and RRC estimators from the linear hazard and ERR models are relatively large in the analysis primarily because of the small event size, as seen in the simulation result. From the three estimators, there is no association between protein intake and breast cancer, which is in general consistent with the literature (Prentice *et al.*, 2009). The association between obesity and breast cancer is not significant from the three estimators, which is likely due to the small sample size with limited cancer events in this data application. From this data application and our simulation study, the results suggest that in the design of a nutritional biomarker study, the event size will likely need to be at least 50 so that the RC and RRC estimators can provide robust measurement error corrections. In the analysis, we applied the Schoenfeld residuals to Cox regression with covariates log(UN percent energy from protein) and obesity, and the proportional-hazards assumption appears to hold (the global test P-value = 0.37). But, to our knowledge model diagnostic for Cox regression with covariate measurement error has not been developed in the literature. From our simulations, analysis based on an additive hazard function is probably not suitable for

TABLE 5 NBS data analysis for time to breast cancer

Cox Regression $r(\beta, X, Z) = \exp(\beta_1 X + \beta_2 Z)$			
	Naive	**RC**	**RRC**
log (% energy from protein /10)			
β	−0.297	1.081	1.007
SE	0.851	2.664	2.664
Obesity			
β	0.313	0.289	0.289
SE	0.484	0.488	0.488
Linear hazard $r(\beta, X, Z) = 1 + \beta_1 X + \beta_2 Z$			
	Naive	**RC**	**RRC**
log (% energy from protein /10)			
β	−0.302	1.956	1.742
SE	0.642	7.971	7.278
Obesity			
β	0.329	0.559	0.538
SE	0.581	1.395	1.307
ERR but without effect modification $r(\beta, X, Z) = (1 + \beta_1 X)$ $\exp(\beta_2 Z)$			
	Naive	**RC**	**RRC**
log (% energy from protein /10)			
β	−0.272	1.701	1.522
SE	0.572	6.614	6.098
Obesity			
β	0.313	0.286	0.288
SE	0.487	0.488	0.488

Note. The estimate and standard errors (SE) for % energy from protein and BMI in the table have been divided by 10 for ease of presentation. The naive estimator replaced the unobserved X by W. The RC estimator replaced X by $E(W|Q, Z)$, and the RRC replaced X by $E(W|Q, Z, T > t)$ at each risk set.

this data set due to the small number of events. The Cox regression model is more appropriate for this data analysis. In this analysis, from our simulation findings, the naive estimator may have underestimated, but the RC and RRC may have overestimated the effect of energy from protein. Future research with more events will likely reduce the finite sample estimation bias from the RC and RRC estimators.

7 | DISCUSSION

In this paper, the semiparametric RC and RRC estimators under a class of general hazard models are investigated for covariate measurement error. Our paper extends the RRC estimator of Xie *et al.* (2001) under Cox regression to the situation when replicates of surrogates may not be available. We also extend the RC and RRC estimators under the ERR model (Wang *et al.*, 2017) to a general class of hazard models. Under exponential hazard regression, the RRC estimator is still inconsistent but it can reduce the bias of the RC estimator. Under linear hazard regression, the RRC estimator is consistent, but the RC estimator has good finite sample performance under many practical situations.

We observed an important finding between the RC and RRC estimators. In measurement error research, the trade-off between bias and efficiency is often referred to the comparison between the naive estimator and a consistent estimator, but it is also applicable to a comparison between RC and RRC in nonlinear regression. In logistic (Liang and Liu, 1991), or Cox regression (Huang and Wang, 2000), RC is biased but it is more efficient than a functional method. However, when the hazard function is linear and the covariate has a symmetric distribution, the RRC estimator not only has smaller biases than the RC estimator but it also has smaller standard errors (although they could be very close). This finding is somewhat different from the general phenomenon of bias-efficiency trade-off that is typically seen in measurement error literature. This is somewhat special, but there could be a possible explanation. The RC estimator under a linear hazard model is theoretically inconsistent, but is somewhat like consistent (see Web Figure S1), which is different from the case in logistic or Cox regression (nonlinear). In contrast, the bias-efficiency trade-off phenomenon does exist in other models because the RC estimator is not consistent under these models.

The semiparametric RC and RRC estimators have a few strengths. First, the calibration function $E(X|Q, \boldsymbol{Z})$ can be implemented by calculating the predicted outcome variable of a working regression model of W given $(Q, \boldsymbol{Z})$. Second, the performance of the RC and RRC estimators is not sensitive to the working model; the RRC estimator is consistent even though the working model is different from the true model. Third, the methods can be applied to the situation when replicates are available. However, the work also has a couple of limitations. First, the IV assumption may not hold in real applications and the RC and RRC estimators may have poor performance if in case the association between Q and X is weak. Second, there are further computing efforts needed to implement the RRC estimator as the calculations are done at each risk set.

ACKNOWLEDGMENTS

This research was partially supported by National Cancer Institute grants CA235122 (Wang), CA239168 (Wang and Song), CA201207 (Song), National Heart, Lung, and Blood Institute grant HL130483 (Wang), NSF grant DMS-1916411 (Song), and a travel award from the Mathematics Research Promotion Center of the Ministry of Science and Technology of Taiwan (Wang).

ORCID

Ching-Yun Wang https://orcid.org/0000-0002-1883-333X

Xiao Song https://orcid.org/0000-0001-8191-7352

REFERENCES

Buonaccorsi, J. (2010) *Measurement Error: Models, Methods, and Applications*. Boca Raton, FL: Hapman and Hall/CRC.

Carroll, R.J., Ruppert, D., Stefanski, L.A. and Crainiceanu, C.M. (2006) *Measurement Error in Nonlinear Models, A modern Perspective*, 2nd edition. London: Chapman and Hall.

Fontana, L., Klein, S. and Holloszy, J.O. (2006) Long-term low-protein, low-calorie diet and endurance exercise modulate metabolic factors associated with cancer risk. *American Journal of Clinical Nutrition*, 84, 1456–1462.

Huang, Y. and Wang, C.Y. (2000) Cox regression with accurate covariates unascertainable: a nonparametric-correction approach. *Journal of the American Statistical Association*, 95, 1209–1219.

Kipnis, V., Midthune, D., Freedman, L. S., Bingham, S., Schatzkin, A., Subar, A. and Carroll, R.J. (2001) Empirical evidence of correlated biases in dietary assessment instruments and its implications. *American Journal of Epidemiology*, 153, 394–403.

Kipnis, V., Subar, A.F., Midthune, D., Freedman, L.S., Ballard-Barbash, R., Troiano, R., Bingham, S., Schoeller, D.A., Schatzkin, A. and Carroll, R.J. (2003) The structure of dietary measurement error: results of the OPEN biomarker study. *American Journal of Epidemiology*, 158, 14–21.

Liang, K.Y. and Liu, X.H. (1991) Estimating equations in generalized linear models with measurement error. In: Godambe, V.P. (Ed.) *Estimating Functions*. Oxford: Clarendon Press, pp 47–63.

Neuhouser, M.L., Tinker, L., Shaw, P.A., Schoeller, D., Bingham, S.A., Horn, L.V., Beresford, S.A.A., Caan, B., Thomson, C., Satterfield, S., Kuller, L., Heiss, G., Smit, E., Sarto, G., Ockene, J., Stefanick, M.L., Assaf, A., Runswick, S. and Prentice, R.L. (2008) Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. *American Journal of Epidemiology*, 167, 1247–1259.

Prentice, R.L. and Mason, M.W. (1986) On the application of linear relative risk regression models. *Biometrics*, 42, 109-120.

Prentice, R.L., Shaw, P.A., Bingham, S.A., Beresford, S.A.A., Caan, B., Neuhouser, M.L., Patterson, R.E., Stefanick, M.L., Satterfield, S., Thomson, C.A., Snetselaar, L., Thomas, A. and Tinker, L. (2009) Biomarker-calibrated energy and protein consumption and increased cancer risk among postmenopausal women. *American Journal of Epidemiology*, 169, 977–989.

Qi, L., Wang, C. Y. and Prentice, R. L. (2005) Weighted estimators for proportional hazards regression with missing covariates. *J. Amer. Statist. Assoc.*, 100, 1250–1263.

Song, X. and Wang, C.Y. (2014) Proportional hazards model with functional covariate measurement error and instrumental variables. *Journal of the American Statistical Association*, 109, 1636–1646.

Sugar, E.A, Wang, C.Y. and Prentice, R.L. (2007) Methods for logistic regression with flexible measurement error. *Biometrics*, 63, 143–151.

Thomas, D.C. (1981) General relative-risk models for survival time and matched case-control analysis. *Biometrics*, 37, 673–686.

Wang, C.Y., Wang, N. and Wang, S. (2000) Regression analysis when covariates are regression parameters of a random effect model for observed longitudinal measurements. *Biometrics*, 56, 487–495.

Wang, C.Y., Cullings, H., Song, X. and Kopecky, K.J. (2017) Joint nonparametric correction estimation for excess relative risk regression in survival analysis. *Journal of the Royal Statistical Society: Series B*, 79, 1583–1599.

Xie, S. X., Wang, C. Y. and Prentice, R. L. (2001) A risk set calibration method for failure time regression using a covariate reliability sample. *Journal of the Royal Statistical Society: Series B*, 63, 855–870.

Wu, L., Hu, X.J. and Wu, H. (2008) Joint inference for nonlinear mixed-effects models and time-to-event at the presence of missing Data. *Biostatistics*, 9, 308–320.

Wu, L., Liu, W. and Hu, J. (2010) Joint inference on HIV viral dynamics and immune suppression in presence of measurement errors. *Biometrics*, 66, 327–335.

SUPPORTING INFORMATION

The PRRC estimator, Web Simulation and Appendix referenced in Sections 1, 4, 5 and 7, along with a zip file for the software, are available with this paper at the Biometrics website on Wiley Online Library.

How to cite this article: Wang C-Y, Song X. Semiparametric regression calibration for general hazard models in survival analysis with covariate measurement error; surprising performance under linear hazard. *Biometrics*. 2021;77:561–572. https://doi.org/10.1111/biom.13318