The International Journal of Biostatistics

Volume 7, Issue 1

2011

Article 4

Regression Calibration with Heteroscedastic Error Variance

Donna Spiegelman, *Harvard School of Public Health*

Roger Logan, *Harvard School of Public Health*

Douglas Grove, *Fred Hutchinson Cancer Research Center*

Recommended Citation:

Spiegelman, Donna; Logan, Roger; and Grove, Douglas (2011) "Regression Calibration with Heteroscedastic Error Variance," *The International Journal of Biostatistics*: Vol. 7 : Iss. 1, Article 4.

Available at: <http://www.bepress.com/ijb/vol7/iss1/4>

DOI: 10.2202/1557-4679.1259

©2011 Berkeley Electronic Press. All rights reserved.

Regression Calibration with Heteroscedastic Error Variance

Donna Spiegelman, Roger Logan, and Douglas Grove

Abstract

The problem of covariate measurement error with heteroscedastic measurement error variance is considered. Standard regression calibration assumes that the measurement error has a homoscedastic measurement error variance. An estimator is proposed to correct regression coefficients for covariate measurement error with heteroscedastic variance. Point and interval estimates are derived. Validation data containing the gold standard must be available. This estimator is a closed-form correction of the uncorrected primary regression coefficients, which may be of logistic or Cox proportional hazards model form, and is closely related to the version of regression calibration developed by Rosner et al. (1990). The primary regression model can include multiple covariates measured without error. The use of these estimators is illustrated in two data sets, one taken from occupational epidemiology (the ACE study) and one taken from nutritional epidemiology (the Nurses' Health Study). In both cases, although there was evidence of moderate heteroscedasticity, there was little difference in estimation or inference using this new procedure compared to standard regression calibration. It is shown theoretically that unless the relative risk is large or measurement error severe, standard regression calibration approximations will typically be adequate, even with moderate heteroscedasticity in the measurement error model variance. In a detailed simulation study, standard regression calibration performed either as well as or better than the new estimator. When the disease is rare and the errors normally distributed, or when measurement error is moderate, standard regression calibration remains the method of choice.

KEYWORDS: measurement error, logistic regression, heteroscedasticity, regression calibration

Author Notes: This study was supported by NIH grants CA50597, NIH ES09411, and NIH CA74112.

1. INTRODUCTION

When validation data are available, the non-iterative *regression calibration* (RC) method can be used to obtain approximately consistent point and interval linear, logistic and Cox regression model parameter estimates with measurement error in one or more continuous covariates, provided certain assumptions are satisfied (Prentice 1982; Fuller 1987; Armstrong, Whittemore et al. 1989; Rosner, Willett et al. 1989; Rosner, Spiegelman et al. 1990; Rosner, Spiegelman et al. 1992; Spiegelman, McDermott et al. 1997; Wang, Hsu et al. 1997; Carroll, Ruppert et al. 2006). In Rosner *et al.*'s version of regression calibration, a standard multiple regression model is used to estimate the uncorrected point and interval estimates of the parameters, and these are bias-corrected using an estimate of the slopes from a linear measurement error model for the true exposure given the surrogate obtained from the validation data. A *Science Citation Index* search in April, 2010 on the three papers by Rosner *et al.* (1989, 1990, 1992) yielded 627 citations, approximately half of which were published in epidemiology and medical journals. Many of these appear to have involved direct applications of the methodology to original analyses of data.

This method, along with others proposed previously for the covariate measurement error problem, including SIMEX (Cook and Stefanski 1995) and the application of empirical process theory to survival data analysis (Huang and Wang 2000), require homoscedastic measurement error variance. The distributions of many environmental and dietary intake variables are often highly skewed, raising concern that the homoscedasticity requirement of regression calibration and other methods may often be unrealistic for important potential applications. For example, in a recent publication, moderate heteroscedasticity was observed in the measurement error model for exposure to airborne soot and nitrogen dioxide (Van Roosbroeck S, Li R et al. 2008), and in a recently published nutritional epidemiology study, moderate heteroscedasticity was observed in the measurement error models for average daily alcohol intake in a pooled analysis of renal cancer incidence (Lee, Hunter et al. 2007). Two other motivating examples of measurement error model heteroscedasticity are studied in depth in this paper, one looking at health symptoms in relation to exposure to anti-neoplastic drugs (Spiegelman and Valanis 1998), and one looking at alcohol intake in relation to breast cancer incidence (Willett, Stampfer et al. 1987). Thus, there is a need to extend regression calibration to apply when the requirement for homoscedasticity of the measurement error model variance is violated, and to compare this extension to several less restrictive iterative approaches, including maximum likelihood and semi-parametric efficient estimating equations (Robins, Hsieh et al. 1995). This paper addresses this need. In Section 2, Rosner *et al.*'s version of regression calibration is reviewed and the new estimator, $\hat{\beta}_{RCH}$, is derived. Next, we consider the case when the true

exposure variable, x , is unobservable, but replicate measures for the unbiased surrogate, X , are available in a reliability sub-study. Then, when heteroscedastic classical measurement error is assumed, we show that it is not possible to obtain a version of the new estimator. Iterative alternatives appropriate to this setting, including maximum likelihood, approximate maximum likelihood, and semi-parametric efficient, are also presented in this section. In Section 3, these estimators are applied to two illustrative examples, one from occupational epidemiology and one from nutritional epidemiology, in which moderate heteroscedasticity is evident. An extensive simulation study of the new estimator, standard regression calibration and the iterative approaches is presented in Section 4. In Section 5, the results of the illustrative examples, of the analytic work and of the simulation study are summarized and recommendations are made.

2. METHODS

The parameter of interest is β_l from the generalized linear model

$$g[E(Y|x, \mathbf{U})] = \beta_0 + \beta_l x + \boldsymbol{\beta}_2^T \mathbf{U}, \quad (1)$$

where Y is the outcome of interest, $g[A]$ is a link function which linearizes the conditional mean function in the covariates and $\mathbf{U}$ is a vector of covariates measured without error. Substituting the covariate measured with error, X , for x , the uncorrected point and interval estimates of effect,

$$\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \hat{\boldsymbol{\beta}}_2^T)^T,$$

are adjusted for measurement error in a one-step procedure. When $g[A] = E(Y|X, \mathbf{U})$, regression calibration is applied to a linear regression model (Fuller 1987). When $g[A] = \text{logit}[E(Y|X, \mathbf{U})]$, regression calibration is applied to a logistic regression model (Rosner, Spiegelman et al. 1990; Rosner, Spiegelman et al. 1992). When

$$\log[I(t)|X, \mathbf{U}] = \log[I(t)|X=0] + \beta_1 X + \boldsymbol{\beta}_2^T \mathbf{U},$$

where $I(t)$ is the incidence rate at time t , then when the disease is rare, regression calibration can be applied to a Cox proportional hazards regression model (Prentice 1982; Spiegelman, McDermott et al. 1997; Wang, Hsu et al. 1997; Hu, Tsiatis et al. 1998). The application of regression calibration to these three basic models, all of which are used widely in epidemiology, was unified with a special focus on interval

estimation and computing in SAS (Spiegelman, McDermott et al. 1997). Application of measurement error methods to data requires that the main study, which contains data $(Y_i, X_i, \mathbf{U}_i), i = 1, \dots, n_1$, is supplemented by an external validation study, which contains data $(X_i, x_i, \mathbf{U}_i), i = n_1 + 1, \dots, n_1 + n_2$. Because it is difficult and expensive to validate X_i with x_i , the validation study is typically much smaller than the main study, i.e. $n_1 \gg n_2$.

The point and interval estimates of effect can be corrected for measurement error using Rosner *et al.*'s formulas (Rosner, Spiegelman et al. 1990)

$$\hat{\boldsymbol{\beta}}_{RC} = \hat{\boldsymbol{\beta}} \hat{\mathbf{F}}^{-1} \quad (2)$$

$$\widehat{Var}(\hat{\boldsymbol{\beta}}_{RC}) \approx \left[\hat{\mathbf{F}}^{-1} \right]^T \widehat{Var}(\hat{\boldsymbol{\beta}}) \hat{\mathbf{F}}^{-1} + \hat{\boldsymbol{\beta}} \widehat{Var}(\hat{\mathbf{F}}^{-1}) \hat{\boldsymbol{\beta}}^T$$

where $\hat{\boldsymbol{\beta}}$ and $\widehat{Var}(\hat{\boldsymbol{\beta}})$ are estimated by fitting (1) to the main study data, $(Y, X, \mathbf{U})$. The first row of $\hat{\mathbf{F}}$, denoted $\hat{\boldsymbol{\gamma}} = (\hat{\gamma}_1, \hat{\gamma}_2^T)^T$, and $\widehat{Var}(\hat{\boldsymbol{\gamma}})$ are obtained from fitting the linear regression model to the validation data

$$E(x | X, \mathbf{U}) = \alpha' + \gamma_1 X + \boldsymbol{\gamma}_2^T \mathbf{U}, \quad (3)$$

under the assumption that

$$Var(x | X, \mathbf{U}) = \sigma^2. \quad (4)$$

Appendix 1 of Rosner, Spiegelman et al. (1990) gives the construction of $\hat{\mathbf{F}}$ from $\hat{\boldsymbol{\gamma}}$ and equation (A7) of the same paper gives $\widehat{Var}(\hat{\mathbf{F}}^{-1})$. Assumption (4) is the homoscedasticity assumption, the relaxation of which is the focus of this manuscript. Regression calibration has been presented by others for use in a variety of settings (Prentice 1982; Fuller 1987; Armstrong, Whittemore et al. 1989; Rosner, Spiegelman et al. 1990; Rosner, Spiegelman et al. 1992; Spiegelman, McDermott et al. 1997; Wang, Hsu et al. 1997; Carroll, Ruppert et al. 2006). All versions of the regression calibration method assume that measurement error is non-differential with respect to the response variable, Y , i.e.

$$f(Y|x, X, \mathbf{U}) = f(Y|x, \mathbf{U}),$$

where $f(\cdot)$ is a density function. In addition to the assumptions specified by (1), (3) and (4), in the case of univariate x , the key additional requirement for approximate unbiasedness of the regression calibration estimator when $g[E(Y|x, \mathbf{U})]$ in (1) is logistic is that either $\beta_1^2 \sigma^2$ is small, or $Pr(Y=1 | x, \mathbf{U})$ is small and $f(x | X, \mathbf{U})$ is normal (Kuha 1994).

It is sometimes the case with biological variables that heteroscedasticity is evident over the observed range of the data; when this occurs, it typically takes the form that the spread increases with the level. Sometimes but by no means always, the variance can be stabilized by log transforming the data but this solution can be undesirable when the variable in question is to be used as a predictor variable in a regression model, and the scientific hypothesis focuses on measuring the relationship of the variable in its originally observed scale to the outcome variable of interest. In this paper, we assume that an empirically verifiable linear model for the mean, (3), holds, but assumption (4), that of constant variance of the measurement error model for $x | X, U$, is untenable for the observed data. Instead, it may appear from the available validation data that

$$Var(x_i | X_i, U_i) = h(X_i, U_i) \sigma^2 \quad (5)$$

is a more reasonable model for the variance, where $h(X_i, U_i)$ is some function of the covariates which induces heteroscedasticity in the measurement error model. In what follows, we derive an extension to the standard regression calibration estimator given above in (2) and to its multivariate counterpart, in which the constant residual variance requirement given by (4) is eliminated.

Standard regression calibration, which assumes homoscedastic measurement error, can be derived by a first order Taylor series expansion around the naïve likelihood in which the mis-measured exposure is treated as if there were no error. In what follows, the second order expansion is developed. By adding the additional term, measurement error model heteroscedasticity can be accommodated, albeit with an unavoidably more complex estimator. In another approach to deriving the regression calibration estimator, the approximate logistic likelihood is derived under the assumptions of normal residual measurement error and rare disease. We develop these approaches in more detail in what follows below, to provide the formal derivation for the new estimator.

We assume the following mean and variance model for x given (X, U) follows

$$\begin{aligned} E(x | X, U) &= \alpha' + \Gamma_1 X + \Gamma_2 U, \\ Var(x | X, U) &= \sigma^2 h(X, U), \end{aligned}$$

$dim(x)=q$, and $h(X, U)$ is a $q \times q$ function. Then, the distribution of $Y | X, U$ for logistic regression with rare disease and multivariate normality is

$$f_1(Y=1 | X, U) \approx e^{\beta_0 + \beta_1^T (\alpha' + \Gamma_1 X + \Gamma_2 U) + \frac{\sigma^2}{2} \beta_1^T h(X, U) \beta_1 + \beta_2^T U} = e^{\beta_0^* + \beta_{11}^{*T} X + \beta_{12}^{*T} h(X, U) \beta_{12}^* + \beta_2^{*T} U} \quad (6)$$

where

$$\begin{aligned}\beta_0^* &= \beta_0 + \beta_1^T \alpha' \\ \beta_{11}^* &= (\beta_1^T \Gamma_1)^T \\ \beta_{12}^* &= \frac{\beta_1 \sigma}{\sqrt{2}} \\ \beta_2^* &= (\beta_1^T \Gamma_2)^T + \beta_2\end{aligned}$$

(Kuha 1994). A similar result for general mean-variance models, using a second-order Taylor series expansion about $E(x|X, U)$ (i.e. a small measurement error approximation) was obtained for scalar x by Carroll *et al.* (Carroll, Ruppert et al. 2006). For $\dim(x) > 1$, Carroll and Stefanski (Carroll and Stefanski 1990) presented an analogous result, but omitted the detailed derivation. Similar approximations have been given for the relative risk function in $\mathbf{X}(t)$, a vector-valued, possibly time-varying covariate, in survival data analysis (Prentice 1982), and for linear regression (Spiegelman, McDermott et al. 1997) where (6) is exact.

Insight can be gained by inspecting the form of (6) with no covariates, $\mathbf{U}$, and for scalar $\mathbf{x}$, denoted x ,

$$f_1(Y | x) = \frac{e^{(\beta_0 + \beta_1 x)Y}}{1 + e^{\beta_0 + \beta_1 x}} \approx e^{(\beta_0 + \beta_1 x)Y}.$$

Figure 1 shows the theoretical relationship between $\log[E(Y)]$ under the rare disease assumption and x given by (1) (solid line), and the approximate induced relationships between $\log[E(Y)]$ and X for $h(X)=X$ (dotted line), $h(X)=X^2$ (short dashed line), and $h(X)=I$, the homoscedastic case, (long dashed lines), when plugging in the measurement error model parameters and logistic regression parameters estimated from the Nurses' Health Study data on breast cancer incidence in relation to alcohol consumption discussed in Section 3.2. It is evident that when $h(X)=X$, the uncorrected estimator is likely to under-estimate β_1 but when $h(X)=X^2$, over-estimation could occur. If $h(X)=X$, (6) simplifies to a linear model in X as a function of the measurement error model and logistic regression model parameters (see Section 2.1), and it can be seen that the standard regression calibration estimator is also likely to over-estimate β_1 (long dashed line).

Figure 1. True, observed and regression calibration-corrected regression functions (NHS)

To simplify notation, $[\beta_{12}^*]^2$ was rewritten as $\tilde{\beta}_{12}^*$. Then, solving for β_1 in the two terms multiplying X and $h(X)$ in (6), two estimators for β_1 ,

$$\hat{\beta}_{11} = \hat{\beta}_{11}^* / \hat{\gamma}_1 \quad (7)$$

and

$$\hat{\beta}_{12} = \text{sign}(\hat{\beta}_{11}) \sqrt{2 |\hat{\beta}_{12}^*| / \hat{\sigma}^2}$$

where $\text{sign}(t) = 1$ if t is positive and -1 if t is negative, are available from the uncorrected regression of Y on X and $h(X)$. Of course, the approximate likelihood (6) can be directly fit to the data using an iterative approach, jointly estimating all parameters simultaneously. Alternatively, we suggest a procedure for obtaining $\hat{\beta}_{RCH,1}$ which can be constructed from standard software tools and used in routine

analysis in what follows. Both approaches will provide consistent estimators, but the maximum likelihood estimator will, of course, be more efficient asymptotically. Later, we will compare the behavior of these estimators in a simulation study.

The new method is as follows:

1. A logistic regression model of Y on X , $h(X, U)$ and U is run in the main study to obtain $\hat{\beta}_{11}^*$ and $\hat{\beta}_{12}^*$ and their estimated variances,
2. A weighted linear regression is run in the validation study, with weights $1/h(X, U)$, to obtain $\hat{\sigma}^2$ and $\hat{\gamma}_1$.
3. $\hat{\beta}_{11}$ and $\hat{\beta}_{12}$ are formed from the formulas above and efficiently combined to produce a single estimate, $\hat{\beta}_{RCH,1}$,

$$\hat{\beta}_{RCH,1} = w_1 \hat{\beta}_{11} + (1 - w_1) \hat{\beta}_{12}$$

The asymptotically minimum variance weights and their derivation, as well as the derivation of the formula for the variance of $\hat{\beta}_{RCH,1}$, are given in Appendix 1. $\hat{\beta}_{RCH,2}$, the measurement-error-corrected estimator of the coefficients corresponding to U , has form

$$\hat{\beta}_{RCH,2} = \hat{\beta}_2^* - \hat{\beta}_{RCH,1} \hat{\gamma}_2 \quad (8)$$

Its asymptotic variance is derived in Appendix 2, along with $Cov(\hat{\beta}_{RCH,1}, \hat{\beta}_{RCH,2})$. As is evident from (6), this estimator can only be used for models with scalar x . For multivariate x with heteroscedastic covariance for $x|X, U$, a term is added to the model for $E(Y|X, U)$ equal to

$$2\sigma^2 \beta_1^T h(X, U; \Sigma) \beta_1,$$

which is a complicated function of the q elements of β_1 that cannot be used to uniquely solve for the q elements of $\hat{\beta}_{12}^*$.

A result similar in form to (6) was given by Prentice (Prentice 1982) for the proportional hazards regression model on a vector $X(t)$, in which one or more of the elements of $X(t)$ are measured with error, and where the conditional distribution of $x(t)$ given $X(t)$ is multivariate normal with a linear mean in $X(t)$ and variance $\Sigma(X(t))$. Hence, under the conditions specified by Prentice, $\hat{\beta}_{RCH}$ can be applied to Cox regression models with an arbitrary number of perfectly measured covariates and a single covariate measured with error, just as discussed above for logistic regression.

Under either small measurement error or a rare disease with normal errors for $x|X, U$, it is evident that if either $\sigma_i^2 = \sigma^2 h(X_i, U_i)$ or β_1^2 is small, for scalar x ,

the standard regression calibration estimator will be approximately valid, even in the presence of heteroscedasticity of the variance of the measurement error model, since the third term in the exponent of (6) vanishes.

2.1 Special case when an ‘alloyed’ gold standard is available

Standard regression calibration is valid even if instead of $\mathbf{x}$, the gold standard, an unbiased imperfect gold standard, $\mathbf{x}_i^*$, is observed in the validation study, where $\mathbf{x}_i^* = \mathbf{x}_i + \varepsilon_i$ and ε_i is a random, mean-zero error term. For example, doubly labeled water is considered an unbiased biomarker for total energy intake (Subar, Kipnis et al. 2003). However, doubly labeled water is measured with some random-within person, unbiased error (Preis, Spiegelman et al. 2010). Here as well, for scalar x , as long as $E(x_i^*) = x_i$ and $Cov(x_i - x_i^*, X_i) = 0$, γ and α' can be consistently estimated from fitting the model

$$E(x_i^* | X_i) = \alpha' + \gamma_1 X_i + \gamma_2 U_i$$

to the data. However,

$$Var(x_i^* | X_i) = h(X, U_i) \sigma^2 + Var(e),$$

where $e_i = x_i - x_i^*$, so Step 2 in the procedure for obtaining $\hat{\beta}_{RCH,1}$ given above will not provide a valid estimate of $h(X_i, U_i) \sigma^2$, the quantity needed for $\hat{\beta}_{RCH,1}$. Without replicate data within subjects, followed by additional calculations not described here, $\hat{\beta}_{RCH}$ is not applicable when an alloyed gold standard is observed.

2.2 Special case of $\hat{\beta}_{RCH}$ when $h(X) = X$

Another important special case emerges when $h(X) = X$. Then, from equation (6) we obtain

$$\hat{\beta}_{RCH,1} = \left[-\hat{\gamma}_1 + \text{sign}(\hat{\beta}_{11}) \sqrt{\hat{\gamma}_1^2 + 2\hat{\sigma}^2 \hat{\beta}_{11}^*} \right] / \hat{\sigma}^2 \quad (9)$$

where $\hat{\beta}_{11}^*$ is the regression coefficient for X obtained from fitting a logistic regression model of Y on X and U , $\hat{\beta}_{11}$ is as above in (7), and $\hat{\beta}_{RCH,2}$ is as above in (8). It is evident from (6) that the function $h(X) = X + b$, where b is a positive constant, can also be considered here. In this case, b is absorbed by the intercept, β_0^* , and does not need to be considered any further in the measurement error correction procedure. The variance of $\hat{\beta}_{RCH,1}$ in the special case was again derived using the multivariate delta method (Bishop, Fienberg et al. 1975). It should be noted that when $\beta_1 \gamma_1 < 0$, $\hat{\beta}_{RCH,1}$ will provide a consistent estimate of β_1 only when

multivariate delta method (Bishop, Fienberg et al. 1975). It should be noted that when $\beta_1 \gamma_1 < 0$, $\hat{\beta}_{RCH,1}$ will provide a consistent estimate of β_1 only when $\frac{1}{2}\sigma^2|\beta_1| < |\gamma_1| < \sigma^2|\beta_1|$ (Appendix 3). However, under (3), as long as $Corr(x, X)$ is positive, as would usually be the case when X is an surrogate for x , γ_1 will be positive, and whenever it is anticipated that β_1 might be negative, x and X can be recoded to avoid this.

2.3 Application to the classical measurement error model

There are an important set of problems where x is unobservable, i.e. no "gold standard" exists. In some of these cases, the measurement error model

$$\begin{aligned} X_{ij} &= x_i + \varepsilon_{ij} \\ E(\varepsilon_{ij}) &= 0 \\ Var(\varepsilon_{ij}) &= \sigma^2 \\ Cov(\varepsilon_{ij}, \varepsilon_{ij'}) &= 0, j \neq j' \\ Cov(\varepsilon_{ij}, x_i) &= 0, j = 1, \dots, n_{2i}, i = 1, \dots, n_2 \end{aligned} \quad (10)$$

is considered reasonable, where n_{2i} is the number of replicates for subject i . Examples of data which are believed to follow this model include blood pressure, serum biomarkers such as cholesterol and its subfractions, hormones, and vitamin concentrations. When model (10) holds, Rosner *et al.*'s (1992) version of the regression calibration method applies, in a procedure similar to that described earlier. In the univariate case, estimates of the reliability coefficient, $Var(x)/Var(X)$, and its variance are substituted for $\hat{\gamma}_1$ and its variance in equation (3). Multivariate generalizations have been given (Rosner, Spiegelman et al. 1992). When the third component of (10) does not hold, i.e. when $Var(\varepsilon_{ij}) = \sigma_i^2 = h(x_i)\sigma^2$, an extension of regression calibration to accommodate this expansion of the model is needed. We rederived the likelihood for scalar x and no additional covariates U to attempt to identify an appropriate estimator in this simple case, now assuming that $X_{ij} | x_i \sim N(x_i, h(x_i)\sigma^2)$, $x_i \sim N(\mu_x, \sigma_x^2)$, $i=1, \dots, n_2$, $j=1, \dots, n_R$, where n_R is the number of replicates for each subject i , and obtained

$$f(x_i, X_{ij}) = \frac{\exp\left\{-\frac{1}{2}\left[\frac{(X_{ij} - x_i)^2}{h(x_i)\sigma^2} + \frac{(x_i - \mu_x)^2}{\sigma_x^2}\right]\right\}}{2\pi\sqrt{h(x_i)\sigma^2\sigma_x^2}},$$

and

$$f(x_i | X_{ij}) = \frac{\exp\left\{-\frac{1}{2}\left[\frac{(X_{ij} - x_i)^2}{h(x_i)\sigma^2} + \frac{(x_i - \mu_x)^2}{\sigma_x^2}\right]\right\}}{2\pi\sqrt{h(x_i)\sigma^2\sigma_x^2}} \cdot \int_{-\infty}^{+\infty} \frac{\exp\left\{-\frac{1}{2}\left[\frac{(X_{ij} - x)^2}{h(x)\sigma^2} + \frac{(x - \mu_x)^2}{\sigma_x^2}\right]\right\}}{2\pi\sqrt{h(x)\sigma^2\sigma_x^2}} dx.$$

Neither $f(x_i | X_{ij})$ nor $f(x_i | X_{ij})$ has a Gaussian structure, and the method of completing the square to obtain a closed-form solution for $f(Y_i | X_i)$ used to derive (6) is therefore not applicable. It is unlikely that a closed-form for

$$f(Y_i | X_i) = \int_x f(Y_i | x) f(x | X_i) dx$$

exists when $Var(\varepsilon_{ij}) = \sigma_i^2 = h(x_i)\sigma^2$ and ε is Gaussian. If functions $h(x_i)$ are found which fit the data at hand, it is unlikely that the resulting expression for $f(Y_i | X_i)$ will be of a form such that the link function, $g[E(Y_i | X_i)]$, can be found with g equal to the logistic or log transformations. Without these features, even $\hat{\beta}_{RCH}$, the scalar version of $\hat{\beta}_{RCH}$, does not exist.

2.4. Iterative methods

Iterative methods can also be applied to the problem of heteroscedastic measurement error in regression covariates. These methods will typically have the advantage of relaxing some of the assumptions required by closed-form methods, such as (1), (3) and (4), but will have the disadvantage of computational complexity which is a barrier to use in applications. In a main study/external validation study design for a binomial outcome and a Gaussian measurement error model, the log likelihood of the data is equal to

$$L(\beta, \alpha', \gamma, \sigma^2) = \sum_{i=1}^{n_1} \log[f_3(Y_i | X_i, U_i; \beta, \alpha', \gamma, \sigma^2)] + \sum_{i=n_1+1}^{n_1+n_2} \log[\phi(x_i | X_i, U_i; \alpha', \gamma, \sigma^2)].$$

We assume here that $\phi(\alpha', \gamma, \sigma_i^2)$ is the normal density with mean given by (3) and variance given by (4), and the probability that $Y_i = 1$ given (X_i, U_i) is

$$\Pr(Y_i | X_i, U_i) = f_3(Y_i | X_i, U_i; \beta, \alpha', \gamma, \sigma^2) = \int_{-\infty}^{\infty} \frac{[\exp(\beta_0 + \beta_1 x + \beta_2^T U_i)]^{Y_i}}{1 + \exp(\beta_0 + \beta_1 x + \beta_2^T U_i)} \frac{\exp\left[-\frac{(x - \alpha' - \gamma_1 X_i - \gamma_2^T U_i)^2}{2h(X_i)\sigma^2}\right]}{\sqrt{2\pi h(X_i)\sigma^2}} dx.$$

The numerical method given by (Crouch and Spiegelman 1990) can be used to evaluate $f_3(Y_i | X_i, U_i; \beta, \alpha', \gamma, \sigma^2)$, or code could be developed which makes use of standard statistical software through the non-linear optimizer that is provided. Generalizing Kuha's derivation (Kuha 1994) to the heteroscedastic measurement error variance case, by using a second-order Taylor series expansion of $\text{logit}[\text{Pr}(Y=1)]$ with respect to β_l around $\beta_l = 0$ under the rare disease assumption and with $x | X, U$ normally distributed with linear mean and variance, $f_3(Y_i | X_i, U_i; \beta, \alpha', \gamma, \sigma^2)$ can be approximated as follows:

$$\begin{aligned} & \text{logit}[f_3(X_i; \beta, \alpha', \gamma, \sigma^2)] \\ & \approx \beta_0 + \beta_l[\alpha' + \gamma^T(X_i, U_i)] + \beta_2^T U_i + \frac{1 - \exp(\beta_0 + \beta_2^T U_i)}{2[1 + \exp(\beta_0 + \beta_2^T U_i)]} \beta_l^2 h(X_i, U_i) \sigma^2. \end{aligned}$$

Likelihoods based on both the exact, $\hat{\beta}_{ML}$, and approximate, $\hat{\beta}_{aML}$, expressions for $f_3(Y_i | X_i, U_i; \beta, \alpha', \gamma, \sigma^2)$ given above were fit to the data in Section 4 and studied by simulation in Section 5.

A consistent, semi-parametric efficient estimator was proposed by (Robins, Hsieh et al. 1995), where (Begun, Hall et al. 1983) defined the semi-parametric efficiency bound as the smallest possible variance obtained by any estimator which is consistent for β over all possible measurement error models for $x | X, U$. Because the model for $x | X, U$ is unknown *a priori* and validation study sample sizes are often small, methods that are robust to mis-specification of the model for $x | X, U$ are desirable. The semi-parametric fully efficient estimator of this class requires a computationally cumbersome non-parametric fit of the density of $x | X, U$ — instead, we investigated the semi-parametric locally efficient version, which is consistent even when the density of $x | X, U$ is mis-specified, and uses a parametric fit for $x | X, U$. As this parametric density, fit and empirically verified in the validation study, approaches the true density for $x | X, U$, full semi-parametric efficiency is approached. Details on this estimator can be found in (Spiegelman and Casella 1997), Appendix 1, where, here $f_2(x | X, U; \alpha', \gamma, \sigma^2)$ was taken to be the normal density with mean given by (3) and variance given by (4), and α', γ , and σ^2 are estimated in the validation study, as usual. This estimator, $\hat{\beta}_{SPLE}$, was fit to the data in Section 4 and studied by simulation in Section 5 along with the others.

3. ILLUSTRATIVE EXAMPLES

3.1 The ACE Study of the acute health effects of occupational exposure to anti-neoplastics among pharmacists

(Valanis, Vollmer et al. 1993) described a cross-sectional study of acute health effects from occupational chemotherapeutics exposure in 675 hospital pharmacists. The research objective is estimation and inference about the prevalence odds ratio for acute health effects related to chemotherapeutics exposure. Here, we will focus on fever prevalence in relation to exposure. There were 110 cases of fever. Average weekly chemotherapeutics exposure (X) was self-reported on questionnaire; in a sub-sample of 56 pharmacists on-site drug mixing diaries were kept for 1-2 weeks (x). The correlation between these two methods of exposure assessment was 0.70. The correlation between the predicted values from the linear measurement error model for diary data (x), conditional upon the questionnaire data (X) and other model covariates, and the absolute value of the residuals (Carroll and Ruppert 1988) was 0.21, indicative of moderate heteroscedasticity (Figure 2).

Figure 2. Evidence for heteroscedasticity in the ACE Study

These analyses were adjusted for three covariates (U): age in years, shift work (1 if night or rotating shift, 0 if day shift), and employed by a community hospital (1 if yes, 0 if no). In previously published analysis of these data, the uncorrected prevalence odds ratio for a top to bottom quintile contrast in number of drugs mixed per day, corresponding to an increment of 52 drugs/day, was 1.08 (95% confidence interval (CI) 1.02-1.15) and the regression calibration estimate of the same quantity, ignoring the observed heteroscedasticity in the measurement error model, was 1.22 (95% CI 1.05- 1.45) (Spiegelman and Valanis 1998). Using maximum likelihood methods with a gamma measurement error model that was empirically verified in the ACE validation study and allows for heteroscedasticity which depends on covariates in an arbitrary manner, the estimated prevalence ratio was 1.17 (95% CI 1.04-1.26) (Spiegelman and Casella 1997). Note that the odds ratio will be a good approximation for the prevalence ratio when the outcome is rare and when the prevalence ratio is near one.

We first needed to identify a form for $h(X)$, and we searched over the class of functions $h(X)=(X+b)^p$. We sought to find the transformation in this class for which the correlation between the absolute value of the weighted residuals from the weighted least squares regression of x on X and the other covariates is nearest to zero, where the weights are $h^{-1}(X)$. As apparent from Table 1, the optimal transformation was $h(X)=X$, leading us to consider the special case of $\hat{\beta}_{RCH}$ discussed in Section 2.2. However, in order to identify the best transformation over the class of functions $h(X)=(X+b)^p$ as shown in Table 1, we needed a non-zero value for b when $p=0$ for the log function, since the data contain zero values. In addition, to fit the SPLE method, a parametric working model for $x|X$ needs to be specified. Here, we used a working normal with mean given by equation (3) and in the footnote of Table 1, and $Var(x|X)=X\sigma^2$. Because the normal likelihood is undefined when $X=0$ and $Var(X)=X\sigma^2$, the function $h(X)=X+b$ was again needed with a small positive value of $b=0.1$. As previously noted, the measurement error correction procedure is invariant to the choice of constant in the function $h(X)=X+b$, and the results in Table 2 were unchanged to three digits of precision for $b=0$.

Table 1. Correlation between the weighted absolute value of the measurement error model residuals and the predicted values for selected weight functions, ACE Study¹

$h(X)$	b	Pearson	Spearman
$(X+b)^0$	0	0.21	0.41
$(X+b)^{0.6}$	0	0.09	0.26
$(X+b)^{0.7}$	0	0.06	0.23
$(X+b)^{0.8}$	0	0.04	0.20
$(X+b)^{0.9}$	0	0.02	0.17
$(X+b)^{1.1}$	0	-0.03	0.10
$(X+b)^{1.4}$	0	-0.09	0.02
$(X+b)^{1.5}$	0	-0.11	-0.02
$(X+b)^{1.7}$	0	-0.15	-0.09
$(X+b)^{1.9}$	0	-0.17	-0.11
$(X+b)^{2.0}$	0	-0.18	-0.12
$\log(X+b)$	0.0001	0.15	0.35
$\log(X+b)$	1.0	0.15	0.35

¹ Model was $E(x) = \alpha' + \gamma^T(X, SHIFT, AGE, COMMHOSP)$, where X is number of drugs mixed per week (questionnaire), $SHIFT$ (1=day shift/0=otherwise), AGE (years), $COMMHOSP$ (1=community hospital/0=otherwise).

Table 2 gives the results of this analysis, where we find that the results which take into account the apparent heteroscedasticity in the measurement error model are virtually unchanged from the standard regression calibration results which ignore this feature of the data. Both the exact and approximate maximum likelihood estimates and the semi-parametric locally efficient estimate were similar to the regression calibration estimates. Application of the semi-parametric estimator led to a large efficiency loss. The regression calibration estimators took trivially more CPU time than the uncorrected estimator, and the iterative estimators took 10 to 100 fold more CPU time than these. With a small data set such as here, none of these CPU times were prohibitive.

Table 2. Comparison of $\hat{\beta}_{RCH}$ to other methods: The ACE Study¹

Method	$\hat{\beta}_l(SE[\hat{\beta}_l]) \times 10^{-3}$	OR (95% CI) ²	p-value	Relative CPU time
Uncorrected	2.37 (0.894)	1.08 (1.02-1.15)	0.008	1 (reference)
$\hat{\beta}_{RC}$	6.01 (2.42)	1.23 (1.04-1.44)	0.013	2
$\hat{\beta}_{RCH}^3$	6.40 (2.59)	1.24 (1.05-1.48)	0.013	2
$\hat{\beta}_{ML}^3$	6.44 (2.68)	1.24 (1.04-1.49)	0.016	35
$\hat{\beta}_{aML}^3$	6.52 (2.68)	1.25 (1.04-1.49)	0.015	119
$\hat{\beta}_{SPL E}$	8.36 (17.5)	1.33 (0.41-4.25)	0.63	356

¹ Adjusted for age, shift, working in a community hospital

² Corresponding to a 34 drug/week increase in mixing activity

³ $h(X)=X+0.1$

3.2 The Nurses' Health Study of the relationship between dietary alcohol intake and breast cancer incidence rates

Willett et al. described a prospective study of the relationship between breast cancer incidence and moderate alcohol consumption among 89,538 U.S. women aged 34-59 who were followed for 4 years beginning in 1980 (Willett, Stampfer et al. 1987). After updating the original data to include 8 years of follow-up, 1466 cases occurred during this study period. Alcohol intake was calculated from three questions about the consumption of beer, wine and liquor that were included on a 61-item food frequency questionnaire data. These data were validated in a sub-sample of 173 women with four one-week weighed diet records (Willett, Sampson et al. 1985). The correlation between these two methods of exposure assessment was 0.85. For average daily alcohol intake (g/day), the correlation between the absolute value of the regression residuals and the predicted values from the linear measurement error model for the diet record data (x) conditional upon the food frequency data (X) and other model covariates was 0.44 (Figure 3).

Figure 3. Evidence for heteroscedasticity in the NHS

A logistic regression model, $\text{logit}(Y_i) = \beta_0 + \beta_1 X_i + \beta_2^T U_i$, was fit to the data, where Y_i is the probability that participant i has received a diagnosis of breast cancer between the time of the 1980 questionnaire return and January 1, 1989, X_i is the covariate measured with error, alcohol intake, and U_i is the vector of other covariates, taken to be perfectly measured: age, age at menarche, menopausal status, age at first live birth, history of benign breast disease, family history of breast cancer, body mass index, and parity. The uncorrected and regression calibration point and interval estimates of the rate ratio from a Cox regression model, corresponding to a 12 g/day increase in alcohol intake, in the Nurses' Health Study based on this same 8 year follow-up period (1466 cases) were given previously as 1.09 (95% CI 1.04-1.14) and 1.15 (95% CI 1.05-1.26), respectively (Spiegelman, McDermott et al. 1997).

Although $h(X) = \log(X + 1.005)$ was an option for minimizing the correlation between the weighted residuals and the predicted from the measurement error model fit in the validation study, transformations of the exposure of interest are not desirable for substantive interpretability unless absolutely necessary for statistical reasons. Thus, the optimal transformation of X for minimizing the correlation between the residuals and the predicted was again a linear one (Table 3), and the special case of $\hat{\beta}_{RCH}$ discussed earlier was again applied.

Table 3. Correlation between the weighted absolute value of the measurement error model residuals and the predicted values for selected weight functions, Nurses' Health Study¹

$h(X)$	b	Outliers included	
		Pearson	Spearman
$(X+b)^0$	0	0.44	0.44
$(X+b)^{0.25}$	0.1	0.40	0.37
$(X+b)^{0.5}$	0.1	0.32	0.26
$(X+b)^{0.5}$	0.5	0.34	0.32
$(X+b)^{0.7}$	0.1	0.22	0.13
$(X+b)^{0.7}$	0.5	0.27	0.24
$(X+b)^{0.9}$	0.1	0.10	-0.01
$(X+b)^{1.15}$	0.1	-0.07	-0.20
$\log(X+b)$	1.0001	-0.22	-0.26
$\log(X+b)$	1.001	-0.14	-0.18
$\log(X+b)$	1.005	0.00	-0.04
$\log(X+b)$	1.008	0.07	0.01
$\log(X+b)$	1.01	0.10	0.03

¹ Model was $E(x) = \alpha' + \gamma^T (X, \text{age}(\text{years}), \text{menopause}, \text{age at menarche} (\leq 12, 13, \geq 14), \text{parity} \times \text{age at first birth (nulliparous/parity 1-2, age at first birth } < 23 / \text{parity 1-2, age at first birth 23-25/parity 1-2, age at first birth 23-25/parity } > 2, \text{ age at first birth } < 23 / \text{parity } > 2, \text{ age at first birth 23-25/parity } > 2, \text{ age at first birth } \geq 26), \text{ history of benign breast disease (no/yes), family history of breast cancer (no/yes), body mass index quintiles 2-5})$

A small non-zero value for the parameter b was used for the reasons given above in Section 3.1. There was a slight attenuation in $\hat{\beta}_{RCH}$ relative to $\hat{\beta}_{RC}$ but substantive interpretation of the data remained unchanged (Table 4). Both the exact and approximate maximum likelihood estimators gave results similar to those given by $\hat{\beta}_{RCH}$, and CPU time was not elevated compared to the regression calibration methods. The semi-parametric efficient estimate was also similar to the others, although slightly larger and less powerful. The excessive CPU time needed for calculations makes this estimator impractical for use with such a large data set.

Table 4. Comparison of $\hat{\beta}_{RCH}$ to other methods: The Nurses' Health Study¹

Method	$\hat{\beta}_l(SE[\hat{\beta}_l]) \times 10^{-3}$	OR (95% CI) ²	p-value	Relative CPU time
Uncorrected	6.96 (2.1)	1.09 (1.04-1.14)	0.00074	1 (reference)
$\hat{\beta}_{RC}$	10.6 (3.2)	1.14 (1.05-1.23)	0.00088	1
$\hat{\beta}_{RCH}^3$	7.50 (2.2)	1.09 (1.04-1.15)	0.00077	1
$\hat{\beta}_{ML}^3$	7.53 (2.2)	1.09 (1.04-1.15)	0.00079	5
$\hat{\beta}_{aML}^3$	7.52 (2.2)	1.09 (1.04-1.15)	0.00077	17
$\hat{\beta}_{SPL}^3$	11.9 (3.4)	1.15 (1.06-1.25)	0.00048	203

¹ Adjusted for age, menopausal status, age at menarche, parity, age at first birth, history of benign breast disease, family history of breast cancer, body mass index quintile

² Corresponding to a 12 g/day increase in alcohol intake

³ $h(X)=X+0.1$

4. SIMULATION STUDY

We studied the small sample behavior of all estimators discussed in Section 2 by simulation. We designed the simulation study to follow the Nurses' Health Study described in Section 3.2, and varied the validation study sample size ($n_2=173$ or 346), the parameter p in $h(X)=X^p$ ($p=0.5, 1, 2$), the extent of measurement error as expressed by $Corr(x, X)=(0.4, 0.6, 0.8)$, and the extent of heteroscedasticity as expressed by $Corr(\hat{e}^2, \hat{x})$ between 0.2 and 0.6. Following the Nurses' Health Study, we set $(\beta_0, \beta_1, \alpha') = (-2.633, 0.01, 3.29)$ and (γ, σ^2) were varied to obtain the desired $Corr(x, X)$ and $Corr(\hat{e}^2, \hat{x})$. To simulate the main study, $n_1=8953$ X 's were chosen with replacement from the Nurses' Health study and used to generate n_1 x 's from a normal distribution with mean given by (3) and variance given by (5) at the specified values of $(\alpha', \gamma, \sigma^2)$. Then, n_1 values of Y were generated from a Bernoulli

with parameter given by (1) with the logistic link function, at the fixed values for (β_0, β_1) . To simulate the validation study, n_2 X 's were chosen with replacement from the Nurses' Health Study validation study and used to generate n_2 x 's from a normal distribution with mean given by (3) and variance given by (5) at the specified values of $(\alpha', \gamma, \sigma^2)$. Five hundred simulations were run for each design point.

Results from the simulation study are given in Table 5. The most striking result is that standard regression calibration, $\hat{\beta}_{RC}$, does as well or better than all the other estimators considered, including maximum likelihood methods, in all scenarios studied. This was true even when measurement error or heteroscedasticity were severe, and from the point of view of both bias and coverage probability. The positive bias predicted in Figure 1 was apparent when $h(X)=X^2$, especially when measurement error was severe. The new estimator, $\hat{\beta}_{RCH}$, performed poorly in many instances, and never did materially better than standard regression calibration. In an attempt to improve finite sample coverage probability, we compared the asymptotic Wald confidence interval coverage probability to the empirical and normalized bootstrapped confidence interval coverage probability, where the latter is the bootstrapped mean of the estimator plus or minus the squared root of the bootstrapped variance times 1.96. Five hundred bootstrapped samples were generated for each of 500 simulations. Bootstrapping the confidence intervals for $\hat{\beta}_{RCH}$ solved the problem of its poor empirical coverage probability. The average bias of the bootstrapped estimator for $\hat{\beta}_{RCH}$ remained larger than that for $\hat{\beta}_{RC}$ in nearly all cases considered, often substantially so, although when bias was calculated as the average of the median bias or median of the median bias, the differences decreased considerably and the overall bias dropped to an acceptable value in many cases (data not shown). In the main and validation study sample sizes considered in this simulation study, the asymptotic optimality of the maximum likelihood estimator and its approximation were not evident.

It is of interest what practical gain is likely be derived from the application of the more robust estimator, $\hat{\beta}_{SPL}$. As can be seen in Table 5, the percent bias, mean square error and coverage probability of $\hat{\beta}_{SPL}$ were acceptable in some of the cases considered, and but no better than those obtained from the standard regression calibration. When measurement error was severe, $\hat{\beta}_{SPL}$ had considerably more bias than standard regression calibration, although its coverage probability was correct. In all cases considered, the computational burden was at least an order of magnitude greater than the maximum likelihood methods and two order of magnitudes greater than the standard regression calibration method.

Table 5. Simulation study of estimators under heteroscedastic measurement error variance

n_2				173						346					
p^I	$\rho_{x,x}$	$\rho_{e^2,x}$		$\hat{\beta}_{UC}$	$\hat{\beta}_{RC}$	$\hat{\beta}_{RCH}$	$\hat{\beta}_{aML}$	$\hat{\beta}_{ML}$	$\hat{\beta}_{SPLE}$	$\hat{\beta}_{UC}$	$\hat{\beta}_{RC}$	$\hat{\beta}_{RCH}$	$\hat{\beta}_{aML}$	$\hat{\beta}_{ML}$	$\hat{\beta}_{SPLE}$
0.5	0.4	0.4	% bias ³	-93	-42	-73	-310	-150	n/a	-91	-12	-32	-110	-80	n/a
			MSE ⁴ $\times 10^{-3}$	0.10	2.7	24	14	4.1		0.096	1.5	26	3.7	2.7	
			Asym CP ⁵ (%)	30	96	66	92	94		29	97	64	95	95	
0.5	0.6	0.4	% bias ³	-100	-110	-180	-143	-134	n/a	-99	-65	-130	-98	-80	n/a
			MSE ⁴ $\times 10^{-3}$	0.12	18	160	18	17		0.11	16	130	16	15	
			Asym CP ⁵ (%)	23	95	82	95	97		22	96	87	96	96	
0.5	0.8	0.5	% bias ³	-81	-14	20	-16	-15	n/a	-79	-3.6	26	-5.8	-5.7	n/a
			MSE ⁴ $\times 10^{-3}$	0.080	0.31	2.0	0.31	0.31		0.075	0.26	2.2	0.26	0.26	
			Asym CP ⁵ (%)	42	94	70	95	95		42	97	67	97	97	
1	0.4	0.5	% bias ³	-80	28	-210	-150	-66	-82	-77	10	-190	-69	-230	-19
			MSE ⁴ $\times 10^{-3}$	0.078	2.5	2.8	1.7	0.74	120	0.072	0.32	2.4	0.84	0.31	0.37
			Asym CP ⁵ (%)	42	97	78	91	95	95	45	97	76	94	96	96
			NormBoot CP ⁶ (%)	0	100	96	n/a	n/a	n/a	0	98	95	n/a	n/a	n/a
			EmpBoot CP ⁷ (%)	0	96	96	n/a	n/a	n/a	0	95	94	n/a	n/a	n/a
			CPU Time	2	2	2	4	13	299	2	2	2	4	13	504
1	0.6	0.5	% bias ³	-61	-0.43	-260	-15	-9.6	n/a	-58	5.4	-200	-3.8	-3.2	n/a
			MSE ⁴ $\times 10^{-3}$	0.052	0.10	4.9	0.19	0.082		0.046	0.078	3.8	0.063	0.065	
			Asym CP ⁵ (%)	62	95	81	95	96		66	95	84	97	97	

Table 5, continued

1	0.8	0.5	% bias ³	-56	-4.0	-630	-6.6	-6.4	4.3	-53	2.0	-460	-0.77	-0.56	n/a
			MSE ⁴ × 10 ⁻³	0.045	0.066	34	0.061	0.062	0.084	0.040	0.054	24	0.050	0.051	
			Asym CP ⁵ (%)	66	95	84	95	96	97	71	96	88	96	96	
			CPU Time	2	2	2	4	12	235	n/a	n/a	n/a	n/a	n/a	
2	0.4	0.6	% bias ³	-60	73	-62	-65	-86	-3427	-59	40	-69	-34	-51	-4929
			MSE ⁴ × 10 ⁻³	0.050	4.8	0.69	1.6	0.62	29	0.048	0.40	0.75	0.17	0.36	1.2
			Asym CP ⁵ (%)	64	95	62	88	88	93	65	96	60	90	91	95
			NormBoot CP ⁶ (%)	0	100	93	n/a	n/a	n/a	0	99	94	n/a	n/a	n/a
			EmpBoot CP ⁷ (%)	0	97	96	n/a	n/a	n/a	0	94	97	n/a	n/a	n/a
			CPU Time	2	2	3	5	14	467	2	2	3	4	14	1005
2	0.6	0.6	% bias ³	6.0	30	2.9	-9.5	-27	n/a	6.7	25	-6.1	-5.3	-5.7	n/a
			MSE ⁴ × 10 ⁻³	0.12	0.054	0.14	0.020	0.50		0.0092	0.034	0.18	0.0050	0.061	
			Asym CP ⁵ (%)	92	89	83	95	95		97	91	91	99	98	
2	0.8	0.6	% bias ³	-55	-2.5	4.2	-7.7	-7.5	n/a	-53	3.8	14	-1.8	-1.2s	n/a
			MSE ⁴ × 10 ⁻³	0.044	0.068	0.17	0.059	0.059		0.039	0.056	0.17	0.048	0.055	
			Asym CP ⁵ (%)	66	94	90	95	95		71	96	91	96	96	

¹ Power in the heteroscedastic variance function $Var(x_i | X_i) = h(X_i)\sigma^2 = X_i^p \sigma^2$; when $p=1$, $h(X_i) = X_i + 0.01$

² $\rho_{x,X} = Corr(x, X)$

³ %bias = $\frac{\sum_{b=1}^{500} \hat{\beta}_b / 500 - \beta}{\beta} \times 100$

⁴ MSE = $\frac{\sum_{b=1}^{500} (\hat{\beta}_b - \beta)^2}{500}$

Table 5, continued

⁵ Asym CP = proportion of simulations in which the asymptotic Wald-based confidence intervals covered the true value of β , where the Wald-based confidence interval is $\hat{\beta}_b \pm 1.96\sqrt{\widehat{Var}_b(\hat{\beta}_b)}$ and $\widehat{Var}_b(\hat{\beta}_b)$ is given by (1).

⁶ NormBoot CP = proportion of simulations in which the normalized bootstrapped confidences intervals covered the true value of β , where the normalized bootstrap confidence interval is $\hat{\beta}_b \pm 1.96\sqrt{\widehat{BootVar}_b(\hat{\beta}_b)}$ and

$$\widehat{BootVar}_b(\hat{\beta}_b) = \frac{\sum_{c=1}^{500} (\hat{\beta}_{bc} - \bar{\beta}_b)^2}{500} \text{ and } \bar{\beta}_b = \frac{\sum_{c=1}^{500} \hat{\beta}_{bc}}{500}$$

⁷ EmpBoot CP = proportion of simulations in which the empirical bootstrapped confidences intervals covered the true value of β , where the b^{th} empirical bootstrap confidence interval are the 5th and 95th percentile values of the set $\hat{\beta}_b = \{\hat{\beta}_{bc}\}, c = 1, 500$

5. DISCUSSION

Although the derivation given in equation (6) has appeared previously, this estimator, $\hat{\beta}_{RCH}$, is novel. The new estimator developed in this paper, $\hat{\beta}_{RCH}$, has several attractive features. Standard software can be used for the primary regression analysis, which is subsequently corrected for bias in a non-iterative calculation at the end. It relaxes one of the most restrictive assumptions of regression calibration, that of homoscedasticity of the measurement error model. This new estimator allows for a multivariate vector of variables measured without error in the primary regression model, along with a scalar variable measured with error, a setting which will be applicable in many situations, including in the examples motivating the research. In two illustrative examples, the new estimator performed well. Although theoretical justification was provided for use of this methodology with the Cox model for rare outcomes and normally distributed error (e.g. (Prentice 1982)), further work could be done to study the behavior of the method in this setting, particularly under departures from rare outcome and from normal errors. In addition, extensions to Poisson regression models with covariate measurement error should also be considered (Fung and Krewski 1999; Kukush, Schneeweis et al. 2004).

It took longer to find $\hat{\beta}_{aML}$ than $\hat{\beta}_{ML}$ in both illustrative examples. Thus, the approximate maximum likelihood estimator should not be considered further.

Although the marginal distributions of x and X were sharply skewed in both the ACE data (for number of drugs mixed per week) and the NHS (for grams of alcohol per day), the distributions of the standardized residuals from the models for $E(x|X)$ were symmetrized to a large extent. Marginal distributions should not, in general, be used as evidence for or against heteroscedasticity in a conditional variance. The correlations between the absolute value of the measurement error model regression residuals and the predicted values from these regressions were moderate. As shown in Section 2, if either β_i^2 or σ^2 is small, the convergent value of $\hat{\beta}_{RCH}$ is likely to be approximately equal to the convergent value of $\hat{\beta}_{RC}$. These conditions appear to have been met in both applications, as in both cases, there was little difference in estimates or inference obtained from the two methods. The approximations used to derive $\hat{\beta}_{RCH}$ assume either rare disease and multivariate normality for the true x given the surrogate, or “small” $\frac{\sigma^2}{2} \beta_i^2 \overline{h(X, U)}$. These assumptions are empirically verifiable, and in the applications in epidemiology which motivated this research, they are verified. In cases where these assumptions are unreasonable, further research is needed to derive suitable estimators.

An extensive simulation study of $\hat{\beta}_{RCH}$ and the other estimators was conducted, based upon a data structure motivated by the Nurses' Health Study data considered in this paper. The results clearly indicated that under the scenarios studied, $\hat{\beta}_{RCH}$ was outperformed by the other estimators. Standard regression calibration, $\hat{\beta}_{RC}$, performed well, as did both the likelihood approximation, $\hat{\beta}_{aML}$, and the exact maximum likelihood estimator, $\hat{\beta}_{ML}$. The good performance of $\hat{\beta}_{RC}$ under heteroscedasticity was observed previously (Spiegelman, Rosner et al. 2000). In the simulation study presented in that paper, two covariates in the logistic regression for Y on $\mathbf{x}$, one with moderate error and the other with considerable error, were dichotomized. By transforming these continuous covariates to Bernoullis, where the variance is a function of the mean, heteroscedasticity was induced. Findings from that study indicated that $\hat{\beta}_{RC}$ was approximately valid for estimation and inference, at least when the validation study size was doubled or more from the original 173. Results from this simulation study suggest that validation study sizes larger than those typically found in nutritional epidemiology are needed when measurement error heteroscedasticity is anticipated.

The coverage probability of $\hat{\beta}_{RCH}$ was below the desired value in nearly all cases considered. This family of heteroscedastic variance functions is similar in spirit to the Box-Cox family of transformations, where, following the initial suggestion of Box and Cox (Box and Cox 1964), standard practice in applied statistics is to estimate the parameter first and then treat this estimate as fixed when estimating the remaining parameters. We did similarly here. This two-step procedure was followed in the computation of the iterative estimators as well. It has been well established that in fitting standard models such as (1) with a heteroscedastic variance function such as (5), the asymptotic distribution of $\hat{\beta}$ is the same whether p in (5) is known or estimated from the data (Carroll and Ruppert 1988). However, in the presence of covariate measurement error, the current situation is somewhat different from the one considered by Carroll and Ruppert, and it is possible that accounting for the estimation of the parameter p , which determines the measurement error variance function within the class considered, could have improved the coverage probabilities. Further research could investigate both the theoretical and empirical properties of the two-step approach in this setting, derive the asymptotic variance of $\hat{\beta}_{RCH}$ with variability of $\hat{p}$ taken into account, examine its variance empirically through simulation studies, and compare to a joint estimation and inference approach for iterative estimators.

$\frac{\sigma^2}{2} \beta_1^2 \overline{h(X,U)}$ ranged over several orders of magnitude, from 0.0000091 to 0.00882 in the simulations shown in Table 5, where $\overline{h(X,U)}$ is the average $h(X,U)$ in the data. Kuha had previously suggested that the bias in $\hat{\beta}_{RC}$ in the homoscedastic measurement error case would be low when $\frac{\sigma^2}{2} \beta_1^2$ was less than 0.5 (Kuha 1994), but we found in another implementation of the regression calibration estimator with homoscedastic measurement error that 0.5 was a far too liberal a criterion, with unacceptable levels of bias when $\frac{\sigma^2}{2} \beta_1^2$ was much smaller than 0.5 (Weller, Milton et al. 2007). In the present simulation study, the Spearman correlation of $\frac{\sigma^2}{2} \beta_1^2 \overline{h(X,U)}$ with bias in $\hat{\beta}_{RC}$ was 0.52 and with bias in $\hat{\beta}_{RCH}$ was 0.22. It appears that $\frac{\sigma^2}{2} \beta_1^2 \overline{h(X,U)}$ is not a reliable metric for identifying conditions under which these estimators may be biased in finite samples in the case of heteroscedastic measurement error.

The maximum likelihood approaches considered here are strictly valid only when the distribution for $x|X,U$ is Gaussian. We did not study the performance of the maximum likelihood estimators when a mis-specified likelihood is fit due to incorrect distributional assumptions about $x|X,U$ or when the proper likelihood is derived under alternative distributions for $x|X,U$. It is indeed likely that the maximum likelihood estimator would exhibit less finite sample bias in a main study/internal validation study design, since this design is substantially more informative than a main study/external validation study design (Spiegelman and Gray 1991). When the disease is rare, as in NHS and, typically, other cohort studies of cancer and other chronic disease endpoints, the external validation study is the design by default. Although the semi-parametric locally efficient estimator, $\hat{\beta}_{SPL}$, is consistent under any distribution for $x|X,U$, this estimator had large finite sample bias in some scenarios studied by simulation, especially as measurement error increased. In addition, this estimator is difficult to program and computationally burdensome to calculate. $\hat{\beta}_{RC}$ can be readily computed using SAS macros obtained at <http://www.hsph.harvard.edu/faculty/spiegelman/blinplus.html> and <http://www.hsph.harvard.edu/faculty/spiegelman/relibpls8.html>.

In summary, as predicted by the theory, standard regression calibration is adequate when measurement error is not severe or the mis-measured covariate effect is moderate, even when heteroscedasticity is severe. It may be worthwhile to recall

that covariate measurement error leads to an exponentially increasing information loss given the same sample size and all other conditions held constant; for example, it is well known that under classical homoscedastic measurement error, the effective sample size is decreased by the factor $\rho_{x,X}^2$ (Fleiss 1986; White, Armstrong et al. 2006). In the most extreme case of measurement error considered in our simulation study, where $\rho_{x,X} = 0.4$, this would lead to more than an 80% reduction in the effective sample size. It appears from the simulation study that heteroscedasticity leads to an even greater information loss for the same amount of measurement error, since $\hat{\beta}_{RC}$ was found to have much better performance in similarly designed simulation studies in prior publications when measurement error was homoscedastic (Rosner, Willett et al. 1989; Carroll and Wand 1991). When measurement error heteroscedasticity is suspected, larger validation studies than are typically the current norm in epidemiology are needed.

Appendix 1. Derivation of the optimal weights and the variance for $\hat{\beta}_{RCH,1}$

Let $V_1 = Var(\hat{\beta}_{11})$, $V_2 = Var(\hat{\beta}_{12})$, and $V_{12} = Cov(\hat{\beta}_{11}, \hat{\beta}_{12})$. By the multivariate delta method,

$$V_1 \approx \frac{Var(\hat{\beta}_{11}^*)}{\gamma_1^2} + \frac{[\hat{\beta}_{11}^*]^2 Var(\hat{\gamma}_1)}{\gamma_1^4}.$$

Under the heteroscedastic measurement error model (5), when γ is estimated using weighted linear regression with weights $h(X_i)$, $i=1, \dots, n_2$,

$$Var(\hat{\gamma}) \approx \sigma^2 [\{ (X_i, U_i^T)^T \}^T diag \{ \frac{1}{h(X_1)}, \frac{1}{h(X_2)}, \dots, \frac{1}{h(X_{n_2})} \} \{ (X_i, U_i^T) \}]^{-1},$$

where $\{ (X_i, U_i^T)^T \}$ is an $n_2 \times (1 + \dim(U))$ matrix with columns (X_i, U_i^T) , $i=1, \dots, n_2$ (Seber 1977). Again by the multivariate delta method,

$$V_2 \approx \frac{Var(\hat{\beta}_{12}^*)}{2\sigma^2 |\beta_{12}^*|} + \frac{|\beta_{12}^*| Var(\hat{\sigma}^2)}{2\sigma^6},$$

since by arguments analogous to those given in Appendix 1 of Spiegelman *et al.* (Spiegelman, Carroll et al. 2001), $Cov(\hat{\beta}_{12}^*, \hat{\sigma}^2)$ is asymptotically 0, and

$$V_{12} \approx sign(\beta_{11}^*) sign(\beta_{12}^*) \frac{Cov(\hat{\beta}_{11}^*, \hat{\beta}_{12}^*)}{\gamma_1 \sigma \sqrt{2 |\beta_{12}^*|}}.$$

All other covariance terms are 0, since by the Gauss-Markov theorem, $Cov(\hat{\gamma}_1, \hat{\sigma}^2) = 0$, and by arguments analogous to those given in Appendix 1 of Spiegelman *et al.* (Spiegelman, Carroll et al. 2001),

$$\text{Cov} \left[\begin{pmatrix} \hat{\beta}_{11}^* \\ \hat{\beta}_{12}^* \end{pmatrix}, \begin{pmatrix} \hat{\gamma}_1 \\ \hat{\sigma}^2 \end{pmatrix} \right] \approx \mathbf{0}.$$

Thus,

$$\text{Var}[w_1 \hat{\beta}_{11} + (1 - w_1) \hat{\beta}_{12}] = w_1^2 V_1 + (1 - w_1)^2 V_2 + 2 w_1 (1 - w_1) V_{12} \quad (1)$$

The estimates of V_1 , V_2 , and V_{12} are obtained by substituting the parameters for their estimates for the uncorrected primary regression of Y on $[X, h(X), \mathbf{U}]$ in the main study, and the weighted linear regression of x on X and $\mathbf{U}$ in the validation study. Likewise, the variances of these parameter estimates are obtained from these same regression analyses, and substituted into the expressions for V_1 , V_2 , and V_{12} to obtain estimates of these quantities.

Now, to derive the optimal weight, w_1 , for (1), we need to minimize

$$g(w_1) = w_1^2 V_1 + (1 - w_1)^2 V_2 + 2 w_1 (1 - w_1) V_{12}$$

with respect to w_1 , since $w_2 = 1 - w_1$, in order to obtain a consistent estimator. The single extremum of this function, subject to the constraint, is at $w_1 = \frac{V_2 - V_{12}}{V_1 + V_2 - 2V_{12}}$.

This is a global minimum as long as $V_1 + V_2 > 2V_{12}$, a condition which is likely to be met in most situations. With these optimal weights, the variance of $\hat{\beta}_{RCH,1}$ is thus

$$\text{Var}(\hat{\beta}_{RCH,1}) = \frac{V_1 V_2 - V_{12}^2}{V_1 + V_2 - 2V_{12}}.$$

To estimate $\text{Var}(\hat{\beta}_{RCH,1})$, estimates of V_1 , V_2 , and V_{12} are obtained from the fit of (1) to the main study data, $\text{Var}(\hat{\gamma}_1)$ is estimated by plugging $\hat{\sigma}^2$ into the expression for $\text{Var}(\hat{\gamma}_1)$ given above, and $\text{Var}(\hat{\sigma}^2) = 2\sigma^4 / [n - \dim(\gamma) - 1]$.

Appendix 2. Derivation of $\text{Var}(\hat{\beta}_{RCH,2})$ and $\text{Cov}(\hat{\beta}_{RCH,1}, \hat{\beta}_{RCH,2})$

By the multivariate delta method,

$$\begin{aligned} \widehat{\text{Var}}(\hat{\beta}_{RCH,2}) &\approx \widehat{\text{Var}}(\hat{\beta}_2^*) + \hat{\beta}_{RCH,1}^2 \widehat{\text{Var}}(\hat{\gamma}_2) + \widehat{\text{Var}}(\hat{\beta}_{RCH,1}) \hat{\gamma}_2 \hat{\gamma}_2^T \\ &\quad - \widehat{\text{Cov}}(\hat{\beta}_2^*, \hat{\beta}_{RCH,1}) \hat{\gamma}_2^T - \hat{\gamma}_2 \widehat{\text{Cov}}(\hat{\gamma}_2, \hat{\beta}_{RCH,1}) \\ &\quad + \hat{\beta}_{RCH,1} \left[\widehat{\text{Cov}}(\hat{\gamma}_2, \hat{\beta}_{RCH,1}) \hat{\gamma}_2^T \right] + \hat{\gamma}_2 \left[\widehat{\text{Cov}}^T(\hat{\gamma}_2, \hat{\beta}_{RCH,1}) \right] \end{aligned}$$

where

$$\widehat{Cov}(\hat{\beta}_2^*, \hat{\beta}_{RCH,1}) \approx \frac{w_1 \widehat{Cov}(\hat{\beta}_2^*, \hat{\beta}_{11}^*)}{\hat{\gamma}_1} + \text{sign}\left(\frac{\hat{\beta}_{11}^*}{\hat{\gamma}_1}\right) \text{sign}(\hat{\beta}_{12}^*) \frac{(1-w_1) \widehat{Cov}(\hat{\beta}_2^*, \hat{\beta}_{12}^*)}{\sqrt{2\hat{\sigma}^2 |\hat{\beta}_{12}^*|}} \quad (2)$$

and

$$\widehat{Cov}(\hat{\gamma}_2, \hat{\beta}_{RCH,1}) \approx -\frac{w_1 \hat{\beta}_{11}^* \widehat{Cov}(\hat{\gamma}_2, \hat{\gamma}_1)}{\hat{\gamma}_1^2} \quad (3)$$

where $\widehat{Var}(\hat{\gamma}_2)$ and $\widehat{Cov}(\hat{\gamma}_2, \hat{\gamma}_1)$ are obtained from the corresponding elements and sub-matrices of

$$Var(\hat{\gamma}) \approx \sigma^2 [\{(X_i, U_i^T)^T\} \text{diag}\{\frac{1}{h(X_1)}, \frac{1}{h(X_2)}, \dots, \frac{1}{h(X_{n_2})}\} \{(X_i, U_i^T)\}^T]^{-1},$$

and $\widehat{Var}(\hat{\beta}_2^*)$, $\widehat{Cov}(\hat{\beta}_2^*, \hat{\beta}_{11}^*)$ and $\widehat{Cov}(\hat{\beta}_2^*, \hat{\beta}_{12}^*)$ are obtained from the uncorrected logistic regression of Y on (X, U) . Finally,

$$\widehat{Cov}(\hat{\beta}_{RCH,2}, \hat{\beta}_{RCH,1}) \approx \widehat{Cov}(\hat{\beta}_2^*, \hat{\beta}_{RCH,1}) - \widehat{Var}(\hat{\beta}_{RCH,1}) \hat{\gamma}_2 - \hat{\beta}_{RCH,1} \widehat{Cov}(\hat{\gamma}_2, \hat{\beta}_{RCH,1})$$

where the covariances in the first and third terms are given by equations (2) and (3), respectively, and $\widehat{Var}(\hat{\beta}_{RCH,1})$ is derived in Appendix 1.

Appendix 3. Proof of conditions under which $\hat{\beta}_{RCH,1}$ is consistent for β_1 when $h(X) = X$

Note that $\beta_{11}^* = \beta_1 \gamma_1 + \beta_1^2 \sigma^2 / 2$ and equation (9) simplifies to

$$\beta_{RCH,1} = \frac{-\gamma_1 + \text{sign}(\beta_{11}) \sqrt{\gamma_1^2 + 2\sigma^2 \beta_{11}^*}}{\sigma^2} = \left[-\gamma_1 + \text{sign}(\beta_1 + \beta_1^2 \sigma^2 / (2\gamma_1)) |\gamma_1 + \beta_1 \sigma^2| \right] / \sigma^2$$

Without loss of generality, hats indicating estimators are suppressed throughout this proof.

$\hat{\beta}_{RCH,1}$ is consistent for β_1 when

$$(1) \quad \gamma_1 + \sigma^2 \beta_1 > 0, \text{ and}$$

$$(2) \quad \beta_1 + \frac{\sigma^2 \beta_1^2}{2\gamma_1} > 0$$

In addition, $\hat{\beta}_{RCH,1}$ is consistent for β_1 when

$$(3) \quad \gamma_1 + \sigma^2 \beta_1 < 0, \text{ and}$$

$$(4) \quad \beta_I + \frac{\sigma^2 \beta_I^2}{2\gamma_I} < 0$$

To see this, first consider the first pair of conditions. When condition 2) holds the sign function yields +1 and condition 1) results in $|\gamma_I + \sigma^2 \beta_I| = \gamma_I + \sigma^2 \beta_I$. Then, the numerator simplifies to $\sigma^2 \beta_I$, and so $\beta_{RCH\ I} = \beta_I$. We now consider the possible combinations of β_I and γ_I ,

I. It is clear that (1) and (2) hold when both β_I and γ_I are positive, hence $\hat{\beta}_{RCH,1}$ is always consistent for β_I under these circumstances.

II. Next, we consider the case when $\gamma_I > 0, \beta_I < 0$.

(1) holds if $-\sigma^2 \beta_I < \gamma_I$, i.e. $\sigma^2 |\beta_I| < \gamma_I$ (*)

(2) holds if $\beta_I + \frac{\sigma^2 \beta_I^2}{2\gamma_I} > 0$ iff $\frac{1}{2} \sigma^2 \beta_I^2 > -\beta_I \gamma_I$

This is the same as $\frac{1}{2} \sigma^2 |\beta_I|^2 > \gamma_I |\beta_I|$. Divide by $|\beta_I|$ to get $\frac{1}{2} \sigma^2 |\beta_I| > \gamma_I$

(**)

(**) $\gamma_I < \frac{1}{2} \sigma^2 |\beta_I|$ and (*) $\sigma^2 |\beta_I| < \gamma_I$ are not possible at the same time. So

both (1) and (2) positive is not possible when $\gamma_I > 0, \beta_I < 0$.

Consider both expressions negative:

3) $\gamma_I + \sigma^2 \beta_I < 0$ iff $\gamma_I < \sigma^2 |\beta_I|$ (*)

4) $\beta_I + \frac{\sigma^2 \beta_I^2}{2\gamma_I} < 0$ iff $\frac{1}{2} \sigma^2 |\beta_I|^2 < -\beta_I \gamma_I = |\beta_I| \gamma_I$

Divide by $|\beta_I|$ to get $\frac{1}{2} \sigma^2 |\beta_I| < \gamma_I$ (**)

$$(*) \gamma_I < \sigma^2 |\beta_I| \text{ and } (**) \frac{1}{2} \sigma^2 |\beta_I| < \gamma_I \text{ yields } \frac{1}{2} \sigma^2 |\beta_I| < \gamma_I < \sigma^2 |\beta_I|.$$

When $\beta_I < 0$ and $\gamma_I > 0$, the only situation that produces a valid estimate is when $\frac{1}{2} \sigma^2 |\beta_I| < \gamma_I < \sigma^2 |\beta_I|$.

III. Next, we consider the case when $\beta_I > 0, \gamma_I < 0$

1) $\gamma_I + \sigma^2 \beta_I > 0$, **and**

$$2) \beta_I + \frac{\sigma^2 \beta_I^2}{2\gamma_I} > 0$$

$$\gamma_I + \sigma^2 \beta_I > 0 \text{ if } \sigma^2 \beta_I > |\gamma_I| \quad (*)$$

$$\beta_I + \frac{\sigma^2 \beta_I^2}{2\gamma_I} > 0 \text{ when } \beta_I > \frac{\sigma^2 \beta_I^2}{2(-\gamma_I)} = \frac{\sigma^2 \beta_I^2}{2|\gamma_I|}. \text{ That is, } |\gamma_I| > \frac{\sigma^2 \beta_I}{2} \quad (**)$$

Both of these are true when $\frac{1}{2} \sigma^2 \beta_I < |\gamma_I| < \sigma^2 \beta_I$.

Now consider, $\hat{\beta}_{RCH,1}$ is consistent for β_I when

3) $\gamma_I + \sigma^2 \beta_I < 0$, **and**

$$4) \beta_I + \frac{\sigma^2 \beta_I^2}{2\gamma_I} < 0$$

$(*) \sigma^2 \beta_I < |\gamma_I|$ and $(**) |\gamma_I| < \frac{1}{2} \sigma^2 \beta_I$ is not possible.

IV) $\beta_I < 0, \gamma_I < 0$

$\hat{\beta}_{RCH,1}$ is consistent for β_I when

1) $\gamma_I + \sigma^2 \beta_I > 0$, **and**

$$2) \beta_I + \frac{\sigma^2 \beta_I^2}{2\gamma_I} > 0$$

It easy to show that both 1) and 2) are never true.

Now consider, $\hat{\beta}_{RCH,1}$ is consistent for β_1 when

$$3) \gamma_1 + \sigma^2 \beta_1 < 0, \text{ and}$$

$$4) \beta_1 + \frac{\sigma^2 \beta_1^2}{2\gamma_1} < 0$$

Similarly it is easy to see that both 3) and 4) are always true. Hence, when $\beta_1 < 0, \gamma_1 < 0$, $\hat{\beta}_{RCH,1}$ is consistent for β_1 .

Conclusion:

$\hat{\beta}_{RCH,1}$ is always consistent for β_1 when both β_1 and γ_1 are positive.

$\hat{\beta}_{RCH,1}$ is always consistent for β_1 when both β_1 and γ_1 are negative.

When $\beta_1 < 0, \gamma_1 > 0$ $\hat{\beta}_{RCH,1}$ is inconsistent for β_1 when $\gamma_1 < \frac{1}{2}\sigma^2 |\beta_1|$ or $\sigma^2 |\beta_1| < \gamma_1$. When $\beta_1 > 0, \gamma_1 < 0$, $\hat{\beta}_{RCH,1}$ is inconsistent for β_1 when $|\gamma_1| < \frac{1}{2}\sigma^2 \beta_1$ or $\sigma^2 \beta_1 < |\gamma_1|$. Hence, when $\beta_1 * \gamma_1 < 0$, $\hat{\beta}_{RCH,1}$ is consistent for β_1 when $\frac{1}{2}\sigma^2 |\beta_1| < |\gamma_1| < \sigma^2 |\beta_1|$.

References

- Armstrong, B. G., A. S. Whittemore, et al. (1989). "Analysis Of Case-Control Data With Covariate Measurement Error - Application To Diet And Colon Cancer." Statistics In Medicine **8**(9): 1151-1163.
- Begun, J. M., W. J. Hall, et al. (1983). "Information and asymptotic efficiency in parametric-nonparametric models." Annals of Statistics **11**: 432-452.
- Bishop, Y. M. M., S. E. Fienberg, et al. (1975). Discrete Multivariate Analyses: Theory and Practice, MIT Press: 492-494.
- Box, G. E. P. and D. R. Cox (1964). "AN ANALYSIS OF TRANSFORMATIONS." Journal of the Royal Statistical Society Series B-Statistical Methodology **26**(2): 211-252.
- Carroll, R. and D. Ruppert (1988). Transformation and weighting in regression. London, Chapman and Hall.

- Carroll, R. J. and D. Ruppert (1988). Transformation and Weighting in Regression. London, Chapman and Hall.
- Carroll, R. J., D. Ruppert, et al. (2006). Measurement Error in Nonlinear Models. London, Chapman & Hall.
- Carroll, R. J. and L. A. Stefanski (1990). "Approximate quasi-likelihood estimation in models with surrogate predictors." Journal of the American Statistical Association **85**: 652-663.
- Carroll, R. J. and M. P. Wand (1991). "SEMIPARAMETRIC ESTIMATION IN LOGISTIC MEASUREMENT ERROR MODELS." Journal of the Royal Statistical Society Series B-Methodological **53**(3): 573-585.
- Cook, J. and L. A. Stefanski (1995). "A simulation extrapolation method for parametric measurement error models." Journal of the American Statistical Association **89**: 1314-1328.
- Crouch, E. A. C. and D. Spiegelman (1990). "The Evaluation Of Integrals Of The Form $\int_0^\infty \int_0^\infty F(T) \exp(-T^2) dt$ - Application To Logistic Normal-Models." Journal Of The American Statistical Association **85**(410): 464-469.
- Fleiss, J. (1986). The design and analysis of clinical experiments. New York, Wiley.
- Fuller, W. A. (1987). Measurement Error Models. New York, Wiley.
- Fung, K. Y. and D. Krewski (1999). "On measurement error adjustment methods in Poisson regression." Environmetrics **10**(2): 213-224.
- Hu, P., A. A. Tsiatis, et al. (1998). "Estimating the parameters in the Cox model when covariate variables are measured with error." Biometrics **54**(4): 1407-1419.
- Huang, Y. and C. Y. Wang (2000). "Cox regression with accurate covariates unascertainable: a nonparametric-correction approach." Journal of the American Statistical Association **95**: 1209-1219.
- Kuha, J. (1994). "Corrections for exposure measurement error in logistic regression models with an application to nutritional data." Stat Med **13**(11): 1135-1148.
- Kukush, A., H. Schneeweis, et al. (2004). "Three estimators for the Poisson regression model with measurement errors." Statistical Papers **45**(3): 351-368.
- Lee, J. E., D. J. Hunter, et al. (2007). "Alcohol intake and renal cell cancer in a pooled analysis of 12 prospective studies." Journal of the National Cancer Institute **99**(10): 801-810.
- Preis, S. R., D. Spiegelman, et al. (2010). "Random and correlated errors in gold standards used in nutritional epidemiology: implications for validation studies." American Journal of Epidemiology **In press**.
- Prentice, R. L. (1982). "Covariate measurement errors and parameter estimation in a failure time regression model." Biometrika **69**: 331-342.

- Robins, J. M., F. Hsieh, et al. (1995). "Semi-parametric efficient estimation of a conditional density with missing or mis-measured covariates." Journal of the Royal Statistical Society, Series B **57**: 409-424.
- Rosner, B., D. Spiegelman, et al. (1990). "Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariates measured with error." Am J Epidemiol **132**(4): 734-745.
- Rosner, B., D. Spiegelman, et al. (1992). "Correction of logistic regression relative risk estimates and confidence intervals for random within-person measurement error." Am J Epidemiol **136**(11): 1400-1413.
- Rosner, B., W. C. Willett, et al. (1989). "Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error." Stat Med **8**(9): 1051-1069; discussion 1071-1053.
- Seber, G. (1977). Linear Regression Analysis. New York, Wiley & Sons.
- Spiegelman, D., R. J. Carroll, et al. (2001). "Efficient regression calibration for logistic regression in main study/internal validation study designs with an imperfect reference instrument." Statistics In Medicine **20**(1): 139-160.
- Spiegelman, D. and M. Casella (1997). "Fully parametric and semi-parametric regression models for common events with covariate measurement error in main study/validation study designs." Biometrics **53**(2): 395-409.
- Spiegelman, D. and R. Gray (1991). "COST-EFFICIENT STUDY DESIGNS FOR BINARY RESPONSE DATA WITH GAUSSIAN COVARIATE MEASUREMENT ERROR." Biometrics **47**(3): 851-869.
- Spiegelman, D., A. McDermott, et al. (1997). "Regression calibration method for correcting measurement-error bias in nutritional epidemiology." Am J Clin Nutr **65**(4 Suppl): 1179S-1186S.
- Spiegelman, D., B. Rosner, et al. (2000). "Estimation and inference for logistic regression with covariate misclassification and measurement error in main study/validation study designs." Journal of the American Statistical Association **95**: 51-61.
- Spiegelman, D. and B. Valanis (1998). "Correcting for bias in relative risk estimates due to exposure measurement error: A case study of occupational exposure to antineoplastics in pharmacists." American Journal of Public Health **88**(3): 406-412.
- Subar, A. F., V. Kipnis, et al. (2003). "Using intake biomarkers to evaluate the extent of dietary misreporting in a large sample of adults: the OPEN study." Am J Epidemiol **158**(1): 1-13.
- Valanis, B. G., W. M. Vollmer, et al. (1993). "Association of antineoplastic drug handling with acute adverse effects in pharmacy personnel." Am J Hosp Pharm **50**: 445-462.
- Van Roosbroeck S, Li R, et al. (2008). "Traffic-related outdoor air pollution and respiratory symptoms in children: the impact of adjustment for exposure measurement error." Epidemiology **19**: 409-416.

- Wang, C. Y., L. Hsu, et al. (1997). "Regression calibration in failure time regression." Biometrics **53**(1): 131-145.
- Weller, E. A., D. K. Milton, et al. (2007). "Regression calibration for logistic regression with multiple surrogates for one exposure." Journal of Statistical Planning and Inference **137**(2): 449-461.
- White, E. J., B. K. Armstrong, et al. (2006). Principles of exposure measurement in epidemiology: collecting, evaluating and improving measures of disease risk factors. Oxford, England; New York, New York, Oxford University Press.
- Willett, W. C., L. Sampson, et al. (1985). "Reproducibility and validity of a semiquantitative food frequency questionnaire." American Journal of Epidemiology **122**: 51-65.
- Willett, W. C., M. J. Stampfer, et al. (1987). "MODERATE ALCOHOL-CONSUMPTION AND THE RISK OF BREAST-CANCER." New England Journal of Medicine **316**(19): 1174-1180.