Journal of Multivariate Analysis 115 (2013) 157–171

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Score tests in the presence of errors in covariates in matched case-control studies

Samiran Sinha *, Seungyoon Yoo

Department of Statistics, Texas A&M University, College Station, TX 77843, United States

ARTICLE INFO

Article history:
Received 20 April 2011
Available online 17 October 2012

AMS subject classifications:
62F03
62J02
62P10

Keywords:
Asymptotic normality
Central limit theorem
Colon cancer
Conditional logistic regression
Score tests
Surrogate measure

ABSTRACT

If covariates are measured with errors, failure to account for that errors may result in a biased estimator of the parameters and consequently the test based on the corresponding estimator may turn out to be biased under the non-zero null hypothesis. In this paper we derive score tests for testing the association between a disease and covariates when a covariate is measured with errors in a matched case-control study. In particular, we deal with the scenario where a possibly biased surrogate is measured in the main data set which is accompanied by an external calibration data that contain the biased surrogate and repeated measures of an unbiased surrogate variable. Under the additive, normal, non-differential measurement errors, and flexible parametric model assumptions, we derive a score test for testing the effect of the covariate measured with errors. In addition, we also derive a score test for a more general hypothesis involving the coefficients associated with the covariates measured with and without errors, which is useful for testing a relationship among the effects of the covariates, such as equality of one or more regression coefficients. Finite sample performance of the proposed method is judged via simulation studies. The proposed method is also applied to a real matched case-control data on colon cancer.

1. Introduction

Errors in a covariate are common in epidemiological studies. In particular, in nutritional epidemiology, association between a disease and nutrient intakes is sought which are usually measured via food frequency questionnaire (FFQ). It is well recognized that nutrient intakes measured via FFQ involve substantial amount of errors [6]. The analysis of the data without reckoning the errors may result in an inconsistent estimator, and consequently any related test may turn out to be biased. In this paper we propose score tests when a covariate is measured with additive errors in a matched case-control study.

Although the naive score test for testing the null hypothesis that there is no association between the disease and the exposure measured with errors is asymptotically valid, usually it loses efficiency [15]. Following the arguments of Tosteson and Tsiatis [15], Carroll et al. [3, Chapter 10] showed that in the generalized linear model, replacing the true exposure by its conditional expectation given the observed covariates and the surrogate in the naive score test, will result in an efficient score test for testing the null that the exposure does not have any effect on the response. Stefanski and Carroll [14] developed a semiparametric score test in the generalized linear model with prospective or cross-sectional data when the exposure is measured with additive errors. In a linear model set-up, Murad and Freedman [9] considered the inference of the interaction term between two covariates measured with additive normal errors. They proposed a method of moment and the regression calibration approach to handle this scenario where both the covariates follow normal distribution. de Castro et al. [4] derived

* Corresponding author.
E-mail address: sinha@stat.tamu.edu (S. Sinha).

0047-259X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2012.10.002

a score test for testing $H_0 : \beta_x = \beta_{x0}$ when covariate X is measured with additive normal errors in a linear model set-up, and β_x denoting the regression coefficient of X in the linear model and β_{x0} could be any non-zero number.

Except de Castro et al. [4], the above mentioned articles mainly focused on the null hypothesis that the disease and the covariate measured with errors are not related. Furthermore, mainly the generalized linear model was used for modeling relation between the response and the covariates based on a prospective or a cross-sectional data. To the best of our knowledge there has been no article which deals with testing of a hypothesis regarding the disease-covariate association in a matched case-control study. One may wonder why we need an additional method for testing a hypothesis in a matched case-control study while there are several tests for prospectively collected data or cross-sectional data. The reason lies in the sampling design of a matched case-control study. First, matched data are clustered by several strata where stratum sizes are bounded whereas the number of strata goes to infinity, in principle. Second, in a matched case-control study, data are collected retrospectively depending on the response variable. Third, there is a stratum-specific parameter in the logistic model relating to the response with the covariates. Finally, the distribution of the covariates may vary across the strata. For handling errors in a covariate in a prospective or cross-sectional data using the structural approach one needs to deal with the conditional distribution of the unobserved true covariate given the error-free covariates whereas in a case-control study one needs to deal with the conditional distribution of the unobserved true covariate given the response variable (disease status) and the other error-free covariates. On the other hand, for a matched case-control data one has to deal with the conditional distribution of the unobserved true covariate given the response variable (disease status), the other error-free covariates, and the matching variables.

In the context of consistent parameter estimation in matched case-control studies in the presence of errors in a covariate, Armstrong et al. [1] proposed a likelihood based approach where they assumed a normal, additive, non-differential measurement errors, and assumed that the erroneous covariate follows a normal distribution. McShane et al. [8] took a more flexible approach for handling normal additive measurement errors that does not require any assumption regarding the distribution of unobserved predictor. In this context, Guolo and Brazzale [5] compared the structural, regression calibration, and the SIMEX approaches through simulation studies only for estimating parameters and not for testing of a hypothesis. To the best of our knowledge there is no paper which considered the measurement errors issue for testing a hypothesis in this set-up.

Our score tests are based on a conditional likelihood function where we use a conditional argument to remove the stratum specific nuisance parameters, and use a flexible normal model for the distribution of the unobserved true covariate. The unobserved stratum specific effect on the distribution of the unobserved true covariate is handled using a random-effect model. Our score tests not only handle test related to the covariate measured with errors but also handle a general hypothesis involving the parameters of the covariates measured with and without errors. This general test is applicable to different scenarios including for a test of equal effect of all the confounding variables. This fact can be illustrated in the context of epidemiological studies of incidence of breast cancer and low-fat diet [10]. For these studies race plays a role of a confounder variable, which has many categories, such as White, Black, Hispanic, American Indian, Asian/Pacific Islander etc. Considering White as a reference category, the log-odds ratio parameter for the disease due to Black, Hispanic, American Indian, Asian/Pacific Islander can be denoted by β_B, β_H, β_{AI}, β_A, respectively. Testing of no difference in the incidence rate of the disease among the racial groups other than the reference group is equivalent to test $H_0 : \beta_B = \beta_H = \beta_{AI} = \beta_A$ which can be tested using our general test procedure.

We would like to point out that in the presence of measurement errors in a covariate, the naive analysis may yield a spurious association between the covariate and the response, and even the relationship between the response and the confounder variables measured without any error may get distorted. Budtz-Jørgensen et al. [2] have nicely illustrated this issue in the context of the linear regression model using a prospective epidemiological study of health effects of prenatal mercury exposure. In their study the response variable was California Verbal Learning Test (CVLT) score, a cognitive task which measures learning and memory. The main covariate was blood mercury concentration which was measured with errors, and location (town) was considered as a confounder. When the errors in the covariate was ignored, CVLT score and location show a strong association. When the errors in the covariate was incorporated in the analysis, the association between the CVLT score and location became statistically insignificant.

A brief outline of the remainder of the paper is as follows. Section 2 contains model, assumption, and a discussion regarding the naive score test. Section 3 contains the proposed methodology. In Section 4 we discuss the situation when the parameters associated with the surrogate variable are estimated from an external calibration data. Section 5 contains simulation studies where we judge and compare the finite sample performance of our proposed approach with that of the naive and the regression calibration approach. The simulation results indicate the advantage and the robustness property of the proposed approach. In Section 6 we apply the proposed method to a real matched case-control data on a colon cancer study. Section 7 concludes with a discussion.

2. Background

Model and assumption: Suppose that we have a $1 : M$ matched case-control data with n strata. Each stratum contains a case (diseased) and M control (non-diseased) subjects. In the data, we observe a set of matching variables S, the binary disease variable Y, a $p \times 1$-vector of error-free covariates Z, and W, an erroneous version of X. We assume that both X and W are scalar variables. Although the extension of the proposed approach to the scenario of multivariate X would follow the same

principles, it may involve tedious algebra and more calculations due to the presence of several integrals in the observed likelihood function. The disease risk model is

$$\text{pr}(Y_{ij} = 1 \mid S_i, X_{ij}, Z_{ij}) = H\{\beta_0(S_i) + \beta_1 X_{ij} + \beta_2^T Z_{ij}\}, \tag{1}$$

where $H(u) = \exp(u)/\{1 + \exp(u)\}$. Here i and j are the indices for strata and the subjects within a stratum. Thus, $i = 1, \ldots, n$, and $j = 1, \ldots, (M + 1)$. Here β_1 and β_2 denote the log-odds ratio parameters corresponding to X and Z, respectively. The effect of the matching variables on the disease risk is conferred through $\beta_0(S_i)$ which is left unspecified. The design of this study implies that $\sum_{j=1}^{M+1} Y_{ij} = 1$ for all i.

When all covariates are error-free, one can test $H_0 : \beta_1 = \beta_{10}$ by using the score statistic

$$\mathcal{T}_0(X, Y, Z; \tilde{\beta}) = S_1(\tilde{\beta})\{I_{n\beta_1\beta_1} - I_{n\beta_1\beta_1} I^{-1}_{n\beta_2\beta_2} I_{n\beta_2\beta_1}\}^{-1} S_1(\tilde{\beta}),$$

where $\tilde{\beta}$ is the MLE of $\beta = (\beta_1, \beta_2^T)^T$ under H_0, i.e., it is obtained by maximizing L_{CLR} under H_0,

$$L_{\text{CLR}} = \prod_{i=1}^{n} \sum_{j=1}^{M+1} Y_{ij} p_{ij}(\beta), \quad p_{ij}(\beta) = \frac{\exp(\beta_1 X_{ij} + \beta_2^T Z_{ij})}{\sum_{k=1}^{M+1} \exp(\beta_1 X_{ik} + \beta_2^T Z_{ik})}.$$

Here, $I_{n\beta_k\beta_l} = -(1/n)[\partial^2 \log(L_{\text{CLR}})/\partial\beta_k \partial\beta_l]_{\beta=\tilde{\beta}}$, $S_1(\beta) = (1/\sqrt{n})\partial \log(L_{\text{CLR}})/\partial\beta_1 = (1/\sqrt{n}) \sum_{i=1}^{n} \sum_{j=1}^{M+1} (Y_{ij} - p_{ij}) X_{ij}$. Under H_0, $\mathcal{T}_0(X, Y, Z; \tilde{\beta})$ follows a χ^2 distribution with 1 degree of freedom. Therefore, we reject H_0 if $\mathcal{T}_0(X, Y, Z; \tilde{\beta}) > \chi^2_{\alpha,1}$, where $\text{pr}(\chi^2_1 > \chi^2_{\alpha,1}) = \alpha$.

We assume that the surrogate variable W (the erroneous measurement of X) is linearly related with X with additive errors

$$W = \delta_0 + \delta_1 X + U_W, \quad \text{where } U_W \sim \text{Normal}(0, \sigma_W^2), \tag{2}$$

and conditional on X, U_W is assumed to be independent of Y (non-differential). If W is an unbiased surrogate for X, then $\delta_0 = 0$ and $\delta_1 = 1$.

Naive score test: The naive score test is $\mathcal{T}_0(W, Y, Z; \tilde{\beta}_{\text{naive}})$, where $\tilde{\beta}_{\text{naive}}$ is the naive estimate of β under H_0. For testing $H_0 : \beta_1 = 0$, the naive score test in the presence of measurement error preserves the level of the test. This fact has been discussed in the context of the generalized linear model, particularly when the observations are a random sample from the population of (Y, W, Z) [15]. However, for testing $H_0 : \beta_1 = \beta_{10} (\neq 0)$ the naive score test may turn out to be a biased test. This fact can be shown mathematically in our context.

Suppose that β_{10} and β_{20} are the true values of β_1 and β_2 under H_0, and $\beta^{\dagger}{}_2$ is the limiting value of the estimated coefficient of Z under H_0 when W is used instead of X. Then the score function used in the naive test can be approximated as

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{M+1} \left\{ Y_{ij} - \frac{\exp(\beta_{10} W_{ij} + \beta^{\dagger T}_2 Z_{ij})}{\sum_{k=1}^{M+1} \exp(\beta_{10} W_{ik} + \beta^{\dagger T}_2 Z_{ik})} \right\} W_{ij} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{M+1} \left\{ Y_{ij} - \frac{\exp(\beta_{10} X_{ij} + \beta^T_{20} Z_{ij})}{\sum_{k=1}^{M+1} \exp(\beta_{10} X_{ik} + \beta^T_{20} Z_{ik})} \right\} W_{ij}$$

$$+ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \sum_{j=1}^{M+1} \left\{ \frac{\exp(\beta_{10} X_{ij} + \beta^T_{20} Z_{ij})}{\sum_{k=1}^{M+1} \exp(\beta_{10} X_{ik} + \beta^T_{20} Z_{ik})} - \frac{\exp(\beta_{10} W_{ij} + \beta^{\dagger T}_2 Z_{ij})}{\sum_{k=1}^{M+1} \exp(\beta_{10} W_{ik} + \beta^{\dagger T}_2 Z_{ik})} \right\} W_{ij}. \tag{3}$$

Since under H_0 the expectation of Y_{ij} conditional on $W_{ij}, X_{ij}, Z_{ij} j = 1, \ldots, (M + 1)$, and $\sum_{j=1}^{M+1} Y_{ij} = 1$ is $\exp(\beta_{10} X_{ij} + \beta^T_{20} Z_{ij}) / \sum_{k=1}^{M+1} \exp(\beta_{10} X_{ik} + \beta^T_{20} Z_{ik})$, the summand of the first term of (3) has expectation zero. Due to the specific structure of the conditional probabilities, when $\beta_{10} \neq 0$, the summand of the second term on the right hand side of (3) has a non-zero expectation irrespective of the dependence between X and Z. As a result, under $H_0 : \beta_1 = \beta_{10} (\neq 0)$, $(1/\sqrt{n}) \sum_{i=1}^{n} \sum_{j=1}^{M+1} \{Y_{ij} - p_{ij}(\beta_{10}, \beta_2^{\dagger})\} W_{ij}$ does not follow an approximate normal with zero mean, making it a biased test. Using a similar argument one can show that the naive test for β_2 is usually biased unless X and Z are independent.

3. Methodology

3.1. Testing the effect of the covariate measured with errors

In the proposed methodology first we assume that $\phi^T = (\delta_0, \delta_1, \sigma_W^2)$ involved in (2) is known. We partition $S = (S_q, S_{nq})$ into two parts, the set of measurable (S_q), and the set of unmeasurable matching variables S_{nq}. The measurable variables

are the ones whose effect can be modeled parametrically, and the effect of S_{nq} cannot be modeled parametrically. For our data example, cases were individually matched with controls based on age and neighborhood of residence. Even though age (S_q) is recorded and measurable, neighborhood would be considered as a unmeasurable (S_{nq}) matching variable. Since the marginal distributions of covariates are not identifiable from the retrospective data, we assume that conditional on the observed covariates Z and S, the unobserved covariate X among the controls follows a normal distribution,

$$[X_{ij} \mid S_i, Z_{ij}, Y_{ij} = 0] \sim \text{Normal}(\alpha_0 + \alpha_1^T S_{iq} + \alpha_2^T Z_{ij} + \gamma_i, \sigma^2). \tag{4}$$

We assume that the mean of X among the controls depends on S_q and Z linearly, and the corresponding association parameters are α_1 and α_2, respectively. The effect of S_{nq} and the nonlinear effect of S_q on the mean of X are captured through γ_i which is assumed to follow a Normal$(0, \sigma_\alpha^2)$. In other words, γ_i can be interpreted as unmeasured stratum specific effect. Indeed one can take a more flexible model for X among controls, and follow our procedure to get a score test. However, the simulation study indicates that normal model among the control population works well for a wide range of scenarios. Now, using models (1) and (4), and following the results of Satten and Kupper [11] and Eq. (6) of Sinha et al. [12] one can show that

$$[X_{ij} \mid S_i, Z_{ij}, Y_{ij} = 1] \sim \text{Normal}(\alpha_0 + \alpha_1^T S_{iq} + \alpha_2^T Z_{ij} + \gamma_i + \beta_1\sigma^2, \sigma^2). \tag{5}$$

From (2), (4) and (5) we obtain the distributions of W_{ij} among the controls and cases $[W_{ij} \mid S_i, Z_{ij}, Y_{ij} = 0] \sim \text{Normal}\{\delta_0 + \delta_1(\alpha_0 + \alpha_1^T S_{iq} + \alpha_2^T Z_{ij} + \gamma_i), \delta_1^2\sigma^2 + \sigma_W^2\}$, $[W_{ij} \mid S_i, Z_{ij}, Y_{ij} = 1] \sim \text{Normal}\{\delta_0 + \delta_1(\alpha_0 + \alpha_1^T S_{iq} + \alpha_2^T Z_{ij} + \gamma_i + \beta_1\sigma^2), \delta_1^2\sigma^2 + \sigma_W^2\}$, which will be used in L_{unobs} given in (7). Following Eq. (6) of Sinha et al. [12] we obtain

$$\frac{\text{pr}(Y_{ij} = 1|S_i, Z_{ij})}{\text{pr}(Y_{ij} = 0|S_i, Z_{ij})} = \exp\left\{\beta_0(S_i) + (\alpha_{0i} + \alpha_1^T S_{iq} + \alpha_2^T Z_{ij})\beta_1 + \beta_1^2\frac{\sigma^2}{2} + \beta_2^T Z_{ij}\right\} \tag{6}$$

which in turn implies

$$\frac{\sum_{j=1}^{M+1} Y_{ij}\text{pr}(Y_{ij} = 1|Z_{ij}, S_i)/\text{pr}(Y_{ij} = 0|Z_{ij}, S_i)}{\sum_{k=1}^{M+1} \text{pr}(Y_{ik} = 1|Z_{ik}, S_i)/\text{pr}(Y_{ik} = 0|Z_{ik}, S_i)} = \sum_{j=1}^{M+1} Y_{ij}P_{ij}, \quad P_{ij} \equiv \frac{\exp\{(\alpha_2^T\beta_1 + \beta_2^T)Z_{ij}\}}{\sum_{k=1}^{M+1}\exp\{(\alpha_2^T\beta_1 + \beta_2^T)Z_{ik}\}}.$$

Conditional on γ_i, the likelihood function given the observed data $(Y_{ij}, W_{ij}, S_i, Z_{ij})$, $j = 1, \ldots, (M+1)$, $i = 1, \ldots, n$ is

$$\begin{aligned} L_{\text{unobs}} &= \prod_{i=1}^{n}\left\{\prod_{j=1}^{M+1} f(W_{ij}|S_i, Z_{ij}, Y_{ij})\right\}\text{pr}\left(Z_{ij}, j = 1, \ldots, (M+1)|S_i, \sum_{j=1}^{M+1} Y_{ij} = 1\right) \\ &= \prod_{i=1}^{n}\left\{\prod_{j=1}^{M+1} f^{Y_{ij}}(W_{ij}|S_i, Z_{ij}, Y_{ij} = 1)f^{1-Y_{ij}}(W_{ij}|S_i, Z_{ij}, Y_{ij} = 0)\right\} \\ &\quad \times \frac{\sum_{j=1}^{M+1} Y_{ij}\text{pr}(Y_{ij} = 1|Z_{ij}, S_i)/\text{pr}(Y_{ij} = 0|Z_{ij}, S_i)}{\sum_{j=1}^{M+1} \text{pr}(Y_{ij} = 1|Z_{ij}, S_i)/\text{pr}(Y_{ij} = 0|Z_{ij}, S_i)}, \end{aligned} \tag{7}$$

where $f(\cdot)$ is the generic symbol for a density function. Observe that although expression (6) is a function of β_{0i} and γ_i along with other parameters, the last component of L_{unobs} does involve neither of the stratum specific parameters due to conditioning on $\sum_{j=1}^{M+1} Y_{ij} = 1$, the sufficient statistic for the stratum specific intercept parameter in the logistic model for Y given Z and S. The observed data likelihood is obtained by integrating out the random effects γ_i from L_{unobs}. That means,

$$L_{\text{obs}} = \int L_{\text{unobs}} \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma_\alpha^2}} \exp\left\{-\frac{\gamma_i^2}{2\sigma_\alpha^2}\right\} d\gamma_i.$$

After plugging in expressions (4)–(6) into L_{unobs}, and then taking the logarithm of L_{obs} we obtain $l = \log(L_{\text{obs}}) = \sum_{i=1}^{n} l_i$, where

$$\begin{aligned} l_i &= -\frac{\sum_{j=1}^{M+1} L_{ij}^2}{2\xi} + \frac{\left(\delta_1\sum_{j=1}^{M+1} L_{ij}\right)^2}{2B\xi^2} - \frac{(M+1)}{2}\log(\xi) - \frac{1}{2}\log(\sigma_\alpha^2) - \frac{1}{2}\log(B) \\ &\quad + \sum_{j=1}^{M+1} Y_{ij}(\alpha_2^T\beta_1 + \beta_2^T)Z_{ij} - \log\left[\sum_{j=1}^{M+1}\exp\{(\alpha_2^T\beta_1 + \beta_2^T)Z_{ij}\}\right], \end{aligned}$$

$L_{ij} \equiv \{W_{ij} - \delta_0 - \delta_1(\alpha_0 + \alpha_1 S_{iq} + \alpha_2 Z_{ij} + \beta_1\sigma^2 Y_{ij})\}$, $\xi \equiv (\delta_1^2\sigma^2 + \sigma_W^2)$, and $B \equiv \{1/\sigma_\alpha^2 + (M+1)\delta_1^2/\xi\}$. Note that all parameters are identifiable as δ_0, δ_1, and σ_W^2 are known here. The score statistic for testing $H_0 : \beta_1 = \beta_{10}$ is

$$\mathcal{T}_1 = S_{\beta_1}(I_{n\beta_1\beta_1} - I_{n\beta_1\theta}I_{n\theta\theta}^{-1}I_{n\theta\beta_1})^{-1}S_{\beta_1},$$

where all components of the statistic are evaluated under H_0, and θ represents all parameters but β_1. Here $S_{\beta_1} \equiv n^{-1/2}\partial l/\partial\beta_1 = S_{\beta_1}(\phi) = \sum_{i=1}^n S_{i,\beta_1}(\phi)$,

$$S_{i,\beta_1}(\phi) = \frac{1}{\sqrt{n}}\left\{\alpha_2^T\sum_{j=1}^{M+1}(Y_{ij} - P_{ij})Z_{ij} + \frac{\delta_1\sigma^2}{\xi}\left(\sum_{j=1}^{M+1}Y_{ij}L_{ij} - \frac{\delta_1^2\sum_{j=1}^{M+1}L_{ij}}{B\xi}\right)\right\},$$

and $I_{n\beta_1\beta_1} \equiv -(1/n)\partial^2 l\partial\beta_1\partial\beta_1 I_{n\beta_1\theta} \equiv -(1/n)\partial^2 l/\partial\beta_1\partial\theta^T$, and $I_{n\theta\theta} \equiv -(1/n)\partial^2 l/\partial\theta\partial\theta^T$. Observe that under H_0, $\beta_1 = \beta_{10}$ and $\theta = \tilde{\theta}$, where $\tilde{\theta}$ satisfies $[(1/n)\partial l/\partial\theta]_{\beta_1=\beta_{10},\theta=\tilde{\theta}} = 0$. The expressions for $\partial l/(\partial\theta)$, $\partial^2 l/(\partial\beta_1\partial\beta_1)$, $\partial^2 l/(\partial\beta_1\partial\theta^T)$ and $\partial^2 l/(\partial\theta\partial\theta^T)$ are given in Appendix A. Under H_0, $\mathcal{T}_1$ follows a χ^2 distribution with 1 degree of freedom.

Remark 1. Based on model assumptions (2) and (4) one may obtain

$$[X_{ij}|S_i, Z_{ij}, W_{ij}, Y_{ij} = 0] \sim \text{Normal}\left[\sigma_\dagger^2\left\{\frac{(W_{ij} + \delta_0)\delta_1}{\sigma_W^2} + (\alpha_0 + \alpha_1^T S_{iq} + \alpha_2^T Z_{ij} + \gamma_i)\sigma^{-2}\right\}, \sigma_\dagger^2\right],$$

where $\sigma_\dagger^2 \equiv (\delta_1^2\sigma_W^{-2} + \sigma^{-2})^{-1}$. Consequently the induced model for the disease odds given S_i, W_{ij}, and Z_{ij} is

$$\begin{aligned}\frac{\text{pr}(Y_{ij} = 1|S_i, W_{ij}, Z_{ij})}{\text{pr}(Y_{ij} = 0|S_i, W_{ij}, Z_{ij})} &= \int \exp(\beta_0(S_i) + \beta_1 X_{ij} + \beta_2^T Z_{ij}) f(X_{ij}|S_i, Z_{ij}, W_{ij}, Y_{ij} = 0)dX_{ij} \\ &= \exp\left[\beta_{0i} + \beta_1\sigma_\dagger^2\left\{\frac{(W_{ij} + \delta_0)\delta_1}{\sigma_W^2} + (\alpha_0 + \alpha_1^T S_{iq} + \alpha_2^T Z_{ij} + \gamma_i)\sigma^{-2}\right\} + \frac{\beta_1^2}{2}\sigma_\dagger^2 + \beta_2^T Z_{ij}\right]\end{aligned}$$

resulting in $\text{pr}(Y_{ij} = 1|S_i, W_{ij}, Z_{ij}) = H\{\beta_0^*(S_i) + \beta_1^* W_{ij} + Z_{ij}\beta_2^*\}$, where $\beta_0^*(S_i) \equiv \beta_0(S_i) + \beta_1^2\sigma_\dagger^2/2 + \beta_1\sigma_\dagger^2\{\delta_0\delta_1\sigma_W^{-2} + (\alpha_0 + \alpha_1^T S_{iq} + \gamma_i)\sigma^{-2}\}$, $\beta_1^* \equiv \beta_1\sigma_\dagger^2\delta_1\sigma_W^{-2}$ and $\beta_2^* = \beta_2 + \beta_1\sigma_\dagger^2\sigma^{-2}\alpha_2$. Therefore, the naive method usually produces a biased estimator for β_1 and β_2. However, when X and Z are independent (i.e., $\alpha_2 = 0$), the naive method yields a consistent estimator for β_2. This fact holds true even when the distribution of X among the controls follows a non-normal model.

3.2. Extension to a more general hypothesis

The above result can be generalized to test $H_0 : Q(\beta) = 0$, where $Q : \mathcal{R}^{(p+1)} \to \mathcal{R}^r$ is a continuous function of $\beta = (\beta_1, \beta_2^T)^T$, and $\partial Q(\beta)/\partial\beta$ is finite for all β in a compact subset of the $(p+1)$-dimensional Euclidean space, and has full row rank r. Define $\psi = (\alpha_0, \alpha_1^T, \alpha_2^T, \sigma_\alpha^2, \sigma^2)^T$.

Result 1. Under the assumptions listed in the Appendix the score test statistic for testing $H_0 : Q(\beta) = 0$ is

$$\begin{aligned}\mathcal{T}_2 = &\begin{pmatrix} n^{-1/2}\sum_{i=1}^n \partial l_i(\tilde{\beta}, \tilde{\psi})/\partial\beta \\ n^{-1/2}\sum_{i=1}^n \partial l_i(\tilde{\beta}, \tilde{\psi})/\partial\psi \end{pmatrix}^T D_n^T(\tilde{\beta}, \tilde{\psi})\hat{A}_{n\beta\cdot\psi}^{-1}\frac{\partial Q(\tilde{\beta})}{\partial\beta}\left[\frac{\partial Q(\tilde{\beta})}{\partial\beta}\hat{A}_{n\beta\cdot\psi}^{-1}D_n(\tilde{\beta}, \tilde{\psi})\tilde{C}\right. \\ &\left.\times D_n^T(\tilde{\beta}, \tilde{\psi})\tilde{A}_{n\beta\cdot\psi}^{-1}\left\{\frac{\partial Q(\tilde{\beta})}{\partial\beta}\right\}^T\right]^{-1}\frac{\partial Q(\tilde{\beta})}{\partial\beta}\tilde{A}_{n\beta\cdot\psi}^{-1}D_n(\tilde{\beta}, \tilde{\psi})\begin{pmatrix} n^{-1/2}\sum_{i=1}^n \partial l_i(\tilde{\beta}, \tilde{\psi})/\partial\beta \\ n^{-1/2}\sum_{i=1}^n \partial l_i(\tilde{\beta}, \tilde{\psi})/\partial\psi \end{pmatrix},\end{aligned}$$

where $D_n(\tilde{\beta}, \tilde{\psi}) \equiv [I - \tilde{A}_{n\beta\psi}\tilde{A}_{n\psi\psi}^{-1}]$, and $\tilde{A}_{n\beta\cdot\psi} \equiv \tilde{A}_{n\beta\beta} - \tilde{A}_{n\beta\psi}\tilde{A}_{n\psi\psi}^{-1}\tilde{A}_{n\psi\beta}$, with $\tilde{A}_{n\beta\beta} = (1/n)[\sum_{i=1}^n \partial^2 l_i/\partial\beta\partial\beta^T]_{\beta=\tilde{\beta},\psi=\tilde{\psi}}$, $\tilde{A}_{n\beta\psi} = (1/n)[\sum_{i=1}^n \partial^2 l_i/\partial\beta\partial\psi^T]_{\beta=\tilde{\beta},\psi=\tilde{\psi}}$, $\tilde{A}_{n\psi\psi} = (1/n)[\sum_{i=1}^n \partial^2 l_i/\partial\psi\partial\psi^T]_{\beta=\tilde{\beta},\psi=\tilde{\psi}}$, and

$$\tilde{C} = \begin{bmatrix} (1/n)\sum_{i=1}^n(\partial l_i/\partial\beta)(\partial l_i/\partial\beta^T) & (1/n)\sum_{i=1}^n(\partial l_i/\partial\beta)(\partial l_i/\partial\psi^T) \\ (1/n)\sum_{i=1}^n(\partial l_i/\partial\psi)(\partial l_i/\partial\beta^T) & (1/n)\sum_{i=1}^n(\partial l_i/\partial\psi)(\partial l_i/\partial\psi^T) \end{bmatrix}_{\beta=\tilde{\beta},\psi=\tilde{\psi}}.$$

Under H_0, $\mathcal{T}_2$ approximately follows a χ_r^2 distribution.

In the proof of this result given in Appendix B we borrowed the techniques given in Theorem 3.5 of White [16]. Note that White [16] considered the generalized score test in a misspecified model. Here we extend that idea in the context of errors-in-covariates in a matched case-control set-up. In the above formulation all components are evaluated at $\beta = \tilde{\beta}$ and $\psi = \tilde{\psi}$, the MLE of β and ψ under H_0. Observe that a special case of this hypothesis is to test $H_0 : \beta_2 = \beta_{20}$ with $Q(\beta) = (0, 1^T)\beta - \beta_{20}$. Similarly, another special case of this hypothesis is to test $H_0 : \beta_1 = \beta_{10}$ with $Q(\beta) = (1, 0^T)\beta - \beta_{10}$.

4. Estimation of δ_0, δ_1, and σ_W^2

Usually the secondary model parameter $\phi = (\delta_0, \delta_1, \sigma_W^2)^T$ is not known, and is estimated from an external calibration data. Following our data example, we consider an external data set consisting of the biased surrogate W and repeated measures of an unbiased surrogate of X. Thus, even in the external data set the true X is not observed, and this scenario is common in nutritional epidemiology. That means the external data contain $(W_l, T_{lk},\ k = 1, \ldots, K \geq 2),\ l = 1, \ldots, m$, where K denotes the number of replicates of T for each of the subject in the calibration data, and m represents the size of the calibration data. In principle, we assume that the calibration sample size $m \to \infty$ along with $n \to \infty$. However, $(n/m) \to \rho \in (0, \infty)$. We assume that

$$T_{lk} = X_l + U_{Tlk}, \quad U_{Tjk} \overset{\text{i.i.d.}}{\sim} \text{Normal}(0, \sigma_T^2).$$

Under the assumption that (1) errors associated with T and W are independent and (2) conditional on X none of U_T and U_W depends on Y, ϕ is consistently estimated by solving the corrected score equations [13]:

$$\sum_{l=1}^{m}(W_l - \widehat{\delta_0} - \widehat{\delta_1}\bar{T}_l) = 0, \quad \sum_{l=1}^{m}(W_l - \widehat{\delta_0} - \widehat{\delta_1}\bar{T}_l)\bar{T}_l + \frac{\widehat{\sigma_T^2}\widehat{\delta_1}}{K} = 0,$$

where $\widehat{\sigma_T^2} = [1/\{m(K-1)\}]\sum_{l=1}^{m}\sum_{k=1}^{K}(T_{lk} - \bar{T}_l)^2$ and $\bar{T}_l = \sum_{k=1}^{K} T_{lk}/K$. Also, $\widehat{\sigma_W^2} = (1/m)\sum_{l=1}^{m}(W_l - \widehat{\delta_0} - \widehat{\delta_1}\bar{T}_l)^2 - \widehat{\delta_1}^2\widehat{\sigma_T^2}/K$. When ϕ is not known, the score statistics given in Sections 3.1 and 3.2 are computed for $\phi = \hat{\phi}$, and consequently the uncertainty of estimation should be taken into account in the score statistics. Observe that $\hat{\phi}$ is a regular linear estimator which let $S_{\beta_1}(\hat{\phi})$ to follow an asymptotic normal distribution. Now using the Taylor series expansion we write $S_{\beta_1}(\hat{\phi}) = S_{\beta_1}(\phi) + \{\partial S_{\beta_1}(\phi)/\partial\phi\}(\hat{\phi} - \phi) + o_p(1)$. Since the external calibration data and the matched case-control data are independently sampled from the population, an estimator of the asymptotic variance of $S_{\beta_1}(\hat{\phi})$ is $\text{var}\{S_{\beta_1}(\phi)\} + \{\partial S_{\beta_1}(\phi)/\partial\phi\}\hat{\Sigma}_\phi\{\partial S_{\beta_1}(\phi)/\partial\phi\}^T$, where $\hat{\Sigma}_\phi$ is the estimated variance of $\hat{\phi}$, and the test statistic $\mathcal{T}_1$ is modified as

$$\mathcal{T}_1^{\text{adj}} = S_{\beta_1}(\hat{\phi})[I_{n\beta_1\beta_1}(\hat{\phi}) - I_{n\beta_1\theta}(\hat{\phi})I_{n\theta\theta}^{-1}(\hat{\phi})I_{n\theta\beta_1}(\hat{\phi}) + \{\partial S_{\beta_1}(\hat{\phi})/\partial\phi\}\hat{\Sigma}_\phi\{\partial S_{\beta_1}(\hat{\phi})/\partial\phi\}^T]^{-1}S_{\beta_1}(\hat{\phi}).$$

Under H_0, $\mathcal{T}_1^{\text{adj}}$ asymptotically follows a χ_1^2 distribution. It is clear that due to the adjustment of the variance of the score, the power of $\mathcal{T}_1^{\text{adj}}$ will be lower than that of $\mathcal{T}_1$. Similarly, when ϕ is estimated from an external data set, $\mathcal{T}_2$ is modified as follows.

Result 2. Under the assumptions listed in the Appendix, and when ϕ is estimated through an external data set, the score test statistic for testing $H_0 : Q(\beta) = 0$ is $\mathcal{T}_2^{\text{adj}}$. Under H_0, $\mathcal{T}_2^{\text{adj}} \overset{\text{approx}}{\sim} \chi_r^2$ where

$$\mathcal{T}_2^{\text{adj}} = \begin{bmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\tilde{\beta}, \tilde{\psi}, \hat{\phi})/\partial\beta \\ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\tilde{\beta}, \tilde{\psi}, \hat{\phi})/\partial\psi \\ \sqrt{\frac{n}{m}}\frac{1}{\sqrt{m}}\sum_{l=1}^{m}\vartheta_l \end{bmatrix}^T D_n(\tilde{\beta}, \tilde{\psi}, \hat{\phi})\hat{A}_{\beta\cdot\psi}^{-1}\left\{\frac{\partial Q(\tilde{\beta})}{\partial\beta}\right\}^T \left[\frac{\partial Q(\tilde{\beta})}{\partial\beta}\hat{A}_{\beta\cdot\psi}^{-1}D_n(\tilde{\beta}, \tilde{\psi}, \hat{\phi})\right.$$

$$\left.\times\tilde{C}D_n(\tilde{\beta}, \tilde{\psi}, \hat{\phi})\hat{A}_{\beta\cdot\psi}^{-1}\left\{\frac{\partial Q(\tilde{\beta})}{\partial\beta}\right\}^T\right]^{-1}\left\{\frac{\partial Q(\tilde{\beta})}{\partial\beta}\right\}\hat{A}_{\beta\cdot\psi}^{-1}D_n^T(\tilde{\beta}, \tilde{\psi}, \hat{\phi})\begin{bmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\tilde{\beta}, \tilde{\psi}, \hat{\phi})/\partial\beta \\ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\tilde{\beta}, \tilde{\psi}, \hat{\phi})/\partial\psi \\ \sqrt{\frac{n}{m}}\frac{1}{\sqrt{m}}\sum_{l=1}^{m}\vartheta_l \end{bmatrix},$$

$$D_n(\tilde{\beta}_n, \tilde{\psi}, \hat{\phi}) \equiv [I \; -\tilde{A}_{n\beta\psi}\tilde{A}^{-1}_{n\psi\psi}\,\tilde{A}_{n\beta\phi} - \tilde{A}_{n\beta\psi}\tilde{A}^{-1}_{n\psi\psi}\tilde{A}_{n\psi\phi}],$$

$$\tilde{C} = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}\{\partial l_i/\partial\beta\}\{\partial l_i/\partial\beta\}^T & \frac{1}{n}\sum_{i=1}^{n}\{\partial l_i/\partial\beta\}\{\partial l_i/\partial\psi^T\} & 0 \\ \frac{1}{n}\sum_{i=1}^{n}\{\partial l_i/\partial\psi\}\{\partial l_i/\partial\beta^T\} & \frac{1}{n}\sum_{i=1}^{n}\{\partial l_i/\partial\psi\}\{\partial l_i/\partial\psi^T\} & 0 \\ 0 & 0 & \left(\frac{n}{m}\right)\frac{1}{m}\sum_{l=1}^{m}\vartheta_l\vartheta_l^T \end{bmatrix}_{\beta=\tilde{\beta},\psi=\tilde{\psi},\phi=\hat{\phi}},$$

$$\vartheta_l = \left\{\frac{1}{m}\sum_{l'=1}^{m}\frac{\partial S_{\phi,l'}}{\partial\phi}\right\}^{-1} S_{\phi,l}, \quad \text{and} \quad S_{\phi,l} = \begin{bmatrix} W_l - \delta_0 - \delta_1\bar{T}_l \\ (W_l - \delta_0 - \delta_1\bar{T}_l)\bar{T}_l + \frac{\sigma_T^2\delta_1}{Km} \\ \sigma_T^2 - \frac{1}{K-1}\sum_{l'=1}^{K}(T_{l'k} - \bar{T}_l)^2 \\ \sigma_W^2 - (W_l - \delta_0 - \delta_1\bar{T}_l)^2 + \frac{\delta_1^2\sigma_T^2}{Km} \end{bmatrix}.$$

The proof of this result is given in Appendix C. Before we conclude this section we would like to reiterate that in our set-up the main surrogate variable W which is observed in the main data set and in the calibration data is possibly biased. However, according to the design of our data example the calibration data contain replicated measurements of an unbiased surrogate variable T along with W. This unbiased surrogate T is usually more expensive to collect. Therefore, it is collected only for a subset of the population.

5. Simulation study

Simulation design: In order to study the performance of the proposed test based on the conditional maximum likelihood (CML) we performed the following simulation study. First, we simulated cohorts of size $N = 50\,000$ by simulating $S = (S_q, S_{nq}), X, Z$, and Y. We took $S_q \sim \text{Normal}(0, 1)$, $S_{nq} \sim \text{Normal}(-0.5, 1)$, $Z \sim \text{Bernoulli}(0.35)$ distribution. We considered two scenarios: (1) $X \sim \text{Normal}(0.5S_q + 0.5S_{nq} + Z, 1)$ and (2) $X \sim 0.25(S_{nq} + S_q + Z) + \text{Gamma}(2, \sqrt{2})$. The second scenario was considered to assess the robustness of the CML method when the model assumption is violated. The binary disease variable Y was simulated from a Bernoulli distribution with the success probability $\text{pr}(Y_{ij} = 1 \mid S_i, X_{ij}, Z_{ij}) = H(\beta_0 + 0.5S_q + 0.5S_{nq} + \beta_1 X + \beta_2 Z)$, and β_0 was varied so that the marginal disease probability varies between 7% and 15%. From the cohort, we constructed 1:2 matched case-control data (i.e., $M = 2$) with $n = 200$ strata. Each stratum consisted of 3 subjects of which one was a case and the rest were controls which were matched with the cases based on S. In order to consider biased surrogate variable we simulated W from (i) $W = 0.5 + X + U_W$ (Table 1) and (ii) $W = 0.5 + 0.75X + U_W$ (Table 2), where $U_W \sim \text{Normal}(0, \sigma_W^2)$. We considered two different values for σ_W^2, $0.25\text{var}(X)$, and $0.5\text{var}(X)$, where $\text{var}(X)$ represents the marginal variance of X. Each matched case-control data were accompanied by a calibration data consisting of (W_l, T_{l1}, T_{l2}), $l = 1, \ldots, m$ with $K = 2$, where $T_{lk} = X_l + U_{T,lk}$, and $U_{T,lk} \sim \text{Normal}(0, \sigma_T^2)$ for $k = 1, 2$, and we set $\sigma_T^2 = 0.5\sigma_W^2$. We took two different values for the size of the calibration data, $m = 25$ and 50. Furthermore, the calibration data were independently sampled from the population.

For each possible combination of the scenarios we tested (a) $H_0 : \beta_1 = 0$ against $H_a : \beta_1 \neq 0$, (b) $H_0 : \beta_1 = 0.5$ against $H_a : \beta_1 \neq 0.5$, and (c) $H_0 : \beta_1 - \beta_2 = 0$ against $H_a : \beta_1 - \beta_2 \neq 0$. For tests (a), (b) and (c) we fixed $\beta_2 = 0.5$ and varied β_1. *Method of analyses*: Under each setting we simulated $R = 2,000$ data sets, and for each simulated data the above tests were conducted using the naive (NV), regression calibration (RC), and the proposed CML approach. For the regression calibration approach we replaced X by $\hat{X} = \hat{\delta}_{0,\text{rc}} + \hat{\delta}_{1,\text{rc}}W$, where $\hat{\delta}_{0,\text{rc}}$ and $\hat{\delta}_{1,\text{rc}}$ were the estimator of the intercept and slope parameters of the regression model relating X with W based on the calibration data. Since the calibration data contain (W, T_1, T_2), $\hat{\delta}_{0,\text{rc}}$ and $\hat{\delta}_{1,\text{rc}}$ are obtained by regressing $(T_1 + T_2)/2$ on W. Consequently, we took into account uncertainties of estimation of $\delta_{0,\text{rc}}$ and $\delta_{1,\text{rc}}$ into the score statistic, and the general form of the adjusted test statistic due to the RC method for testing $H_0 : Q(\beta) = 0$ is given by $T^{\text{adj}}_{2,\text{rc}}$. In deriving this formula we used somewhat analogous steps as of Result 2.

$$T^{\text{adj}}_{2,\text{rc}} = \begin{bmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_{i,\text{rc}}(\tilde{\beta}_{\text{rc}}, \hat{\delta}_{\text{rc}})/\partial\beta \\ \sqrt{\frac{n}{m}}\frac{1}{\sqrt{m}}\sum_{l=1}^{m}\vartheta_{l,\text{rc}} \end{bmatrix}^T D_{n,\text{rc}}(\tilde{\beta}_{\text{rc}}, \hat{\delta}_{\text{rc}})\hat{A}^{-1}_{n\beta\beta,\text{rc}}\left\{\frac{\partial Q(\tilde{\beta}_{\text{rc}})}{\partial\beta}\right\}^T \left[\frac{\partial Q(\tilde{\beta}_{\text{rc}})}{\partial\beta}\hat{A}^{-1}_{n\beta\beta,\text{rc}}\right.$$

$$\left.\times D_{n,\text{rc}}(\tilde{\beta}_{\text{rc}}, \hat{\delta}_{\text{rc}})\tilde{C}_{\text{rc}}D_{n,\text{rc}}(\tilde{\beta}_{\text{rc}}, \hat{\delta}_{\text{rc}})\hat{A}^{-1}_{n\beta\beta,\text{rc}}\left\{\frac{\partial Q(\tilde{\beta}_{\text{rc}})}{\partial\beta}\right\}^T\right]^{-1}\left[\left\{\frac{\partial Q(\tilde{\beta}_{\text{rc}})}{\partial\beta}\right\}\hat{A}^{-1}_{n\beta\beta,\text{rc}}D_{n,\text{rc}}(\tilde{\beta}_{\text{rc}}, \hat{\delta}_{\text{rc}})\right]$$

Table 1
Results of the simulation study for comparing powers of several tests. Here, NV, RC, and CML represent the score tests based on the naive, the regression calibration, and the proposed conditional maximum likelihood method, respectively. For all simulations we fix $\beta_2 = 0.5$, and $W = 0.5 + X + U$, with $U \sim \text{Normal}(0, \sigma_W^2)$, $\kappa \equiv \sigma_W^2/\sigma_X^2$ and $\sigma_T^2 = 0.5\sigma_W^2$.

κ	0	NV		$m = 25$				$m = 50$			
		0.25	0.5	RC		CML		RC		CML	
				0.25	0.5	0.25	0.5	0.25	0.5	0.25	0.5
β_1	$X \sim \text{Normal}(0.5S_{nq} + 0.5S_q + Z, 1)$										
	$H_0 : \beta_1 = 0$ versus $H_0 : \beta_1 \neq 0$										
−0.2	0.641	0.467	0.382	0.449	0.332	0.327	0.158	0.492	0.421	0.429	0.205
−0.1	0.23	0.182	0.15	0.171	0.116	0.115	0.067	0.148	0.116	0.116	0.106
0	0.045	0.044	0.045	0.037	0.029	0.026	0.007	0.045	0.045	0.023	0.011
0.1	0.224	0.169	0.145	0.154	0.109	0.127	0.056	0.182	0.140	0.159	0.112
0.2	0.634	0.485	0.388	0.453	0.33	0.332	0.145	0.444	0.333	0.409	0.224
	$H_0 : \beta_1 = 0.5$ versus $H_0 : \beta_1 \neq 0.5$										
0.3	0.594	0.369	0.577	0.271	0.358	0.244	0.212	0.258	0.352	0.292	0.266
0.4	0.198	0.296	0.555	0.200	0.326	0.094	0.085	0.193	0.296	0.121	0.083
0.5	0.045	0.232	0.527	0.133	0.248	0.033	0.038	0.126	0.249	0.045	0.048
0.6	0.154	0.176	0.496	0.095	0.214	0.120	0.094	0.088	0.201	0.133	0.111
0.7	0.47	0.142	0.468	0.077	0.195	0.296	0.194	0.067	0.167	0.343	0.243
	$H_0 : \beta_1 - \beta_2 = 0$ versus $H_0 : \beta_1 - \beta_2 \neq 0$										
0.3	0.129	0.38	0.562	0.272	0.351	0.117	0.098	0.267	0.355	0.119	0.108
0.4	0.068	0.318	0.561	0.201	0.311	0.065	0.070	0.194	0.297	0.077	0.078
0.5	0.052	0.246	0.535	0.13	0.240	0.059	0.061	0.133	0.241	0.056	0.054
0.6	0.059	0.191	0.484	0.087	0.197	0.072	0.071	0.095	0.199	0.075	0.083
0.7	0.120	0.134	0.464	0.073	0.163	0.104	0.097	0.071	0.170	0.114	0.101
	$X \sim 0.25(S_{nq} + S_q + Z) + \text{Gamma}(2, \sqrt{2})$										
	$H_0 : \beta_1 = 0$ versus $H_0 : \beta_1 \neq 0$										
−0.2	0.575	0.535	0.440	0.450	0.354	0.243	0.202	0.502	0.424	0.292	0.258
−0.1	0.159	0.127	0.120	0.160	0.133	0.140	0.127	0.190	0.170	0.176	0.154
0	0.056	0.051	0.048	0.050	0.049	0.046	0.045	0.049	0.046	0.050	0.044
0.1	0.218	0.171	0.155	0.162	0.155	0.160	0.142	0.201	0.188	0.198	0.141
0.2	0.698	0.642	0.532	0.551	0.412	0.342	0.286	0.567	0.435	0.389	0.341
	$H_0 : \beta_1 = 0.5$ versus $H_0 : \beta_1 \neq 0.5$										
0.3	0.624	0.264	0.366	0.157	0.171	0.129	0.112	0.159	0.167	0.288	0.188
0.4	0.209	0.163	0.276	0.086	0.095	0.065	0.071	0.091	0.101	0.152	0.095
0.5	0.047	0.067	0.145	0.080	0.092	0.032	0.028	0.076	0.079	0.034	0.028
0.6	0.180	0.064	0.096	0.052	0.066	0.126	0.082	0.059	0.055	0.149	0.102
0.7	0.496	0.042	0.073	0.126	0.114	0.284	0.178	0.135	0.116	0.352	0.215
	$H_0 : \beta_1 - \beta_2 = 0$ versus $H_0 : \beta_1 - \beta_2 \neq 0$										
0.3	0.156	0.271	0.364	0.153	0.166	0.090	0.070	0.163	0.165	0.099	0.078
0.4	0.085	0.165	0.279	0.088	0.098	0.063	0.055	0.092	0.093	0.065	0.058
0.5	0.054	0.085	0.172	0.086	0.090	0.040	0.051	0.067	0.077	0.040	0.031
0.6	0.080	0.054	0.102	0.067	0.085	0.046	0.052	0.055	0.058	0.056	0.058
0.7	0.138	0.051	0.068	0.136	0.144	0.069	0.063	0.132	0.121	0.089	0.072

$$\times \begin{bmatrix} \dfrac{1}{\sqrt{n}} \displaystyle\sum_{i=1}^{n} \partial l_{i,\text{rc}}(\tilde{\beta}_{\text{rc}}, \hat{\delta}_{\text{rc}})/\partial\beta \\ \sqrt{\dfrac{n}{m}} \dfrac{1}{\sqrt{m}} \displaystyle\sum_{l=1}^{m} \vartheta_{l,\text{rc}} \end{bmatrix},$$

where $D_{n,\text{rc}}(\tilde{\beta}_{\text{rc}}, \hat{\delta}_{\text{rc}}) = [I\ \hat{A}_{n\beta\beta,\text{rc}}]$, $\hat{A}_{n\beta\beta,\text{rc}} = (1/n)[\sum_{i=1}^{n} \partial^2 l_{i,\text{rc}}/\partial\beta\partial\beta^T]_{\beta=\tilde{\beta}_{\text{rc}}}$, $l_{i,\text{rc}} = \log \sum_{j=1}^{M+1} Y_{ij} p_{ij,\text{rc}}$, $p_{ij,\text{rc}} = \exp(\beta_{1,\text{rc}}\hat{X}_{ij} + \beta_{2,\text{rc}}^T Z_{ij}) / \sum_{k=1}^{M+1} \exp(\beta_{1,\text{rc}}\hat{X}_{ik} + \beta_{2,\text{rc}}^T Z_{ik})$,

$$\tilde{C}_{\text{rc}} = \begin{bmatrix} \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} (\partial l_{i,\text{rc}}/\partial\beta)(\partial l_{i,\text{rc}}/\partial\beta)^T & 0 \\ 0 & \left(\dfrac{n}{m}\right) \dfrac{1}{m} \displaystyle\sum_{l=1}^{m} \vartheta_{l,\text{rc}} \vartheta_{l,\text{rc}}^T \end{bmatrix}_{\beta=\tilde{\beta}_{\text{rc}}, \delta=\hat{\delta}_{\text{rc}}}$$

Table 2
Results of the simulation study for comparing powers of several tests. Here, NV, RC, and CML represent the score tests based on the naive, the regression calibration, and the proposed conditional maximum likelihood method, respectively. For all simulations we fix $\beta_2 = 0.5$, and $W = 0.5 + 0.75X + U$, with $U \sim \text{Normal}(0, \sigma_W^2)$, $\kappa \equiv \sigma_W^2/\sigma_X^2$, and $\sigma_T^2 = 0.5\sigma_W^2$.

κ	0	NV		$m = 25$				$m = 50$			
		0.25	0.5	RC		CML		RC		CML	
				0.25	0.5	0.25	0.5	0.25	0.5	0.25	0.5
β_1	$X \sim \text{Normal}(0.5S_{nq} + 0.5S_q + Z, 1)$										
	$H_0 : \beta_1 = 0$ versus $H_0 : \beta_1 \neq 0$										
−0.2	0.63	0.408	0.313	0.406	0.300	0.194	0.100	0.426	0.309	0.272	0.133
−0.1	0.193	0.154	0.121	0.156	0.117	0.068	0.052	0.134	0.122	0.076	0.065
0	0.048	0.049	0.050	0.046	0.048	0.023	0.002	0.048	0.053	0.017	0.006
0.1	0.202	0.141	0.107	0.139	0.108	0.076	0.062	0.148	0.112	0.082	0.067
0.2	0.625	0.411	0.298	0.402	0.284	0.201	0.112	0.385	0.311	0.282	0.142
	$H_0 : \beta_1 = 0.5$ versus $H_0 : \beta_1 \neq 0.5$										
0.3	0.601	0.852	0.993	0.610	0.589	0.252	0.142	0.638	0.622	0.276	0.188
0.4	0.202	0.557	0.956	0.362	0.412	0.115	0.090	0.352	0.402	0.124	0.102
0.5	0.054	0.286	0.848	0.165	0.254	0.042	0.032	0.146	0.223	0.048	0.037
0.6	0.147	0.086	0.646	0.103	0.157	0.109	0.072	0.076	0.113	0.115	0.082
0.7	0.451	0.054	0.454	0.144	0.138	0.231	0.109	0.112	0.074	0.264	0.161
	$H_0 : \beta_1 - \beta_2 = 0$ versus $H_0 : \beta_1 - \beta_2 \neq 0$										
0.3	0.129	0.397	0.646	0.338	0.433	0.086	0.076	0.338	0.458	0.094	0.079
0.4	0.068	0.361	0.679	0.287	0.408	0.064	0.059	0.277	0.412	0.073	0.078
0.5	0.052	0.296	0.675	0.213	0.370	0.063	0.055	0.219	0.379	0.053	0.053
0.6	0.059	0.221	0.653	0.174	0.351	0.073	0.075	0.168	0.343	0.077	0.077
0.7	0.120	0.177	0.671	0.133	0.305	0.099	0.080	0.142	0.328	0.103	0.083
	$X \sim 0.25(S_{nq} + S_q + Z) + \text{Gamma}(2, \sqrt{2})$										
	$H_0 : \beta_1 = 0$ versus $H_0 : \beta_1 \neq 0$										
−0.2	0.575	0.423	0.321	0.430	0.329	0.193	0.176	0.404	0.305	0.227	0.185
−0.1	0.159	0.133	0.113	0.143	0.115	0.103	0.076	0.152	0.127	0.120	0.101
0	0.056	0.048	0.054	0.049	0.051	0.043	0.038	0.049	0.050	0.048	0.045
0.1	0.218	0.174	0.144	0.163	0.129	0.112	0.120	0.169	0.146	0.150	0.123
0.2	0.698	0.539	0.445	0.507	0.401	0.208	0.180	0.500	0.411	0.245	0.220
	$H_0 : \beta_1 = 0.5$ versus $H_0 : \beta_1 \neq 0.5$										
0.3	0.646	0.646	0.923	0.414	0.360	0.212	0.111	0.413	0.356	0.259	0.142
0.4	0.219	0.296	0.714	0.214	0.191	0.121	0.058	0.198	0.162	0.132	0.068
0.5	0.051	0.085	0.436	0.172	0.178	0.029	0.022	0.167	0.126	0.033	0.029
0.6	0.176	0.059	0.159	0.120	0.155	0.094	0.069	0.077	0.092	0.119	0.072
0.7	0.491	0.225	0.067	0.391	0.291	0.205	0.109	0.326	0.279	0.252	0.137
	$H_0 : \beta_1 - \beta_2 = 0$ versus $H_0 : \beta_1 - \beta_2 \neq 0$										
0.3	0.156	0.215	0.341	0.162	0.165	0.039	0.030	0.165	0.157	0.054	0.033
0.4	0.085	0.113	0.264	0.091	0.107	0.029	0.031	0.092	0.097	0.032	0.031
0.5	0.054	0.065	0.161	0.076	0.098	0.027	0.019	0.072	0.081	0.022	0.022
0.6	0.080	0.051	0.089	0.062	0.082	0.033	0.036	0.056	0.074	0.041	0.039
0.7	0.138	0.076	0.061	0.135	0.112	0.057	0.065	0.144	0.121	0.072	0.069

and

$$\vartheta_{l,\text{rc}} = \begin{bmatrix} m & \sum_{l=1}^{m} W_l \\ \sum_{l=1}^{m} W_l & \sum_{l=1}^{m} W_l^2 \end{bmatrix}^{-1} \begin{bmatrix} (\bar{T}_l - \hat{\delta}_{0,\text{rc}} - \hat{\delta}_{1,\text{rc}} W_l) \\ (\bar{T}_l - \hat{\delta}_{0,\text{rc}} - \hat{\delta}_{1,\text{rc}} W_l) W_l \end{bmatrix}.$$

For the CML method δ_0, δ_1, and σ_W^2 were estimated from the calibration data using the method described in Section 5. Thus, for testing the hypotheses we used $\mathcal{T}_1^{\text{adj}}$ and $\mathcal{T}_2^{\text{adj}}$. For the sake of comparison we also present the power of the test $\mathcal{T}_0$ when there are no measurement errors.

Simulation results: For testing $H_0 : \beta_1 = 0$, the NV, RC, and the CML method maintain the level for all scenarios. However, the NV shows more power than the RC and the CML approach. For testing $H_0 : \beta_1 = 0.5$, the NV and the RC method do not maintain the nominal levels, and result in biased tests whereas the CML method maintains the level and the power

of the CML method increases as β_1 moves away from 0.5. Furthermore, the power of the CML approach increases with the size of the calibration data. Similarly, for testing $H_0 : \beta_1 - \beta_2 = 0$, the NV and the RC method turn out to be biased. Overall, as $\kappa = \sigma_W^2/\sigma_X^2$ increases the power of the CML method decreases. The simulation results indicate that the proposed method is quite robust for moderate departure from the normal distribution assumption of X. The results also indicate that overall the power of the CML method is lower when $E(W|X) = 0.5 + 0.75X$ compared to the scenario when $E(W|X) = 0.5 + X$.

6. Analysis of the colon cancer data

Background information: Here we analyze a 1:1 matched case-control data on colon cancer which was taken from the Study of Diet and Health conducted in two metropolitan areas Toronto and Calgary of Canada during the period 1976–1978 [7]. The matched data consisted of $n = 171$ male colon cancer patients aged between 35 and 75, and each case was matched with a control based on age and neighborhood of residence. We refer to Jain et al. [7] for details on recruiting cases and controls into the study. Apart from demographic information each subject responded to a diet questionnaire which was aimed at measuring nutrient intakes. This data set was previously analyzed by Armstrong et al. [1].

In this study we will be looking for the effect of dietary fiber on colon cancer. Since fiber intake measured via dietary questionnaire involves measurement errors, an external calibration study was conducted to estimate this error. In the external study $m = 16$ healthy volunteers reported dietary histories considered as W along with the detailed weighted food records kept by the volunteers' spouses. The later variable is considered as an unbiased measurement (T) for the true dietary intakes.

Analyses and results: In the analysis we consider body mass index as Z, and log of dietary fiber intake as X, Y denotes the disease status as if a subject has colon cancer or not. The log of dietary fiber measured via diet questionnaire is considered as W. The assumed model is $\text{pr}(Y = 1|S, X, Z) = H(\beta_0(S) + \beta_1 X + \beta_2 Z)$, and we test $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$ at 5% level of significance, where β_1 represents the log-odds ratio parameter corresponding to dietary fiber intake. For the naive test we obtain $\mathcal{T}_0(W, Y, Z, \tilde{\beta}_{\text{Naive}}) = 0.462$. Thus, we do not have sufficient evidence to reject H_0 at 5% level. Next, we compute the score test based on the RC approach. It may be noted here that we first estimate $E(X|W) = \delta_{0,\text{rc}} + \delta_{1,\text{rc}} W$ based on the calibration data. The test statistic due to the RC method is 0.353. Hence, we do not reject H_0 at 5% level of significance. For the CML approach we get $\hat{\delta_0} = 3.16$, $\hat{\delta_1} = 0.303$, $\hat{\sigma_W^2} = 0.0499$, and the test statistic is $\mathcal{T}_1^{\text{adj}} = 0.266$. Hence, we also do not reject H_0 at 5% level. Since, H_0 is accepted based on all three approaches, there is no need to test any non-zero value for β_1. However, for the purpose of illustration we conduct a test if the odds ratio corresponding to log of fiber intake is 2. That means we test $H_0 : \beta_1 = \log(2)$ against $H_1 : \beta_1 \neq \log(2)$. The test statistics due to the naive, regression calibration, and the proposed methods are 12.562, 5.272, and 4.889, respectively, and the corresponding theoretical p-values are 0.0004, 0.022, and 0.027, respectively. We also conduct a test for the log-odds ratio parameter corresponding to log of BMI. Here we test $H_0 : \beta_2 = \log(2)$ against $H_1 : \beta_2 \neq \log(2)$. The test statistics due to the naive, regression calibration, and the proposed methods are 0.012, 0.011, and 0.044, respectively, and the corresponding theoretical p-values are 0.913, 0.916, and 0.834, respectively. Therefore, we fail to reject the null hypothesis here.

7. Discussion

We propose score tests for testing hypothesis in a matched case-control study when a covariate is measured with errors which in turn may help for finding optimal sample size for designing such studies. The methods include a general score test involving several model parameters. We also provide the theory when the errors are calibrated from an external data set, and incorporated into the analysis. Although, in the proposed method we assumed a normal model for X among controls conditional on the stratification variables, and the error-free covariates, the simulation study indicates very satisfactory performance of the method even when the model assumption is moderately violated. One of the limitations of the proposed method and any other methods which correct bias due to measurement errors is inflated uncertainty resulting in large standard error of estimators or low power of any test. This uncertainty does not decrease by increasing sample size.

In principle, the proposed approach can be applied to when one takes a more flexible model for X and the error distribution. In future, we will develop score tests for handling multiplicative measurement errors which are also common in observational studies. Another potential research problem is to develop a generalized score test when both the distributions of X and the errors are unspecified.

Acknowledgments

The authors thank the referee for useful suggestions which have led to a much improved version of the manuscript. The authors wish to thank Mr. P.G. Ghosh for his careful reading of the manuscript and suggestions to improve English. This research was partially supported by NSF grant SES-0961618.

Appendix A. Components of the score vector and the Hessian matrix

Define $\theta \equiv (\beta_2^T, \alpha_0, \alpha_1^T, \alpha_2^T, \sigma_\alpha^2, \sigma^2)^T$, thus $(\partial l/\partial\theta) = [(\partial l/\partial\beta_2)^T, (\partial l/\partial\alpha_0), (\partial l/\partial\alpha_1)^T, (\partial l/\partial\alpha_2)^T, (\partial l/\partial\sigma_\alpha^2), (\partial l/\partial\sigma^2)]^T$. Also, define $\zeta = (\delta_1/\xi)\{1 - \delta_1^2/(B\xi)\}$. Then the score functions are

$$\frac{\partial l}{\partial\beta_1} = \alpha_2 \sum_{i=1}^{n}\sum_{j=1}^{M+1}(Y_{ij} - p_{ij})Z_{ij} + \frac{\delta_1\sigma^2}{\xi}\sum_{i=1}^{n}\sum_{j=1}^{M+1}L_{ij}\left(Y_{ij} - \frac{\delta_1^2}{\xi}\right), \qquad \frac{\partial l}{\partial\beta_2} = \sum_{i=1}^{n}\sum_{j=1}^{M+1}(Y_{ij} - p_{ij})Z_{ij},$$

$$\frac{\partial l}{\partial\alpha_0} = \zeta\sum_{i=1}^{n}\sum_{j=1}^{M+1}L_{ij}, \frac{\partial l}{\partial\alpha_1} = \zeta\sum_{i=1}^{n}\left(S_{qi}\sum_{j=1}^{M+1}L_{ij}\right), \qquad \frac{\partial l}{\partial\alpha_2} = \zeta\sum_{i=1}^{n}\sum_{j=1}^{M+1}L_{ij}Z_{ij} + \beta_1\sum_{i=1}^{n}\sum_{j=1}^{M+1}(Y_{ij} - p_{ij})Z_{ij},$$

$$\frac{\partial l}{\partial\sigma_\alpha^2} = -\frac{n}{2\sigma_\alpha^2} + \frac{n}{2B\sigma_\alpha^4} + \frac{\delta_1^2}{2B^2\sigma_\alpha^4\xi^2}\sum_{i=1}^{n}\left(\sum_{j=1}^{M+1}L_{ij}\right)^2,$$

$$\begin{aligned}\frac{\partial l}{\partial\sigma^2} &= \frac{\delta_1^2}{\xi^2}\sum_{i=1}^{n}\left(0.5\sum_{j=1}^{M+1}L_{ij}^2 - \frac{\beta_1\delta_1}{B}\sum_{j=1}^{M+1}L_{ij}\right) + \frac{\beta_1\delta_1}{\xi}\sum_{i=1}^{n}\sum_{j=1}^{M+1}L_{ij}Y_{ij}\\ &\quad + \frac{\delta_1^4}{2B\xi^3}\left\{\frac{\delta_1^2(M+1)}{B\xi} - 2\right\}\sum_{i=1}^{n}\left(\sum_{j=1}^{M+1}L_{ij}\right)^2 - \frac{n(M+1)\delta_1^2}{2\xi B}\left(\frac{1}{\sigma_\alpha^2} + \frac{M\delta_1^2}{\xi}\right).\end{aligned}$$

The components of the Hessian matrix are

$$\frac{\partial^2 l}{\partial\beta_1^2} = -\alpha_2^T\sum_{i=1}^{n}\left\{\sum_{j=1}^{M+1}Z_{ij}Z_{ij}^T p_{ij} - \left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)\left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)^T\right\}\alpha_2 - \frac{n\delta_1^2\sigma^4}{\xi B}\left(\frac{1}{\sigma_\alpha^2} + \frac{M\delta_1^2}{\xi}\right),$$

$$\frac{\partial^2 l}{\partial\beta_2^T\partial\beta_1} = -\alpha_2^T\sum_{i=1}^{n}\left\{\sum_{j=1}^{M+1}Z_{ij}Z_{ij}^T p_{ij} - \left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)\left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)^T\right\}, \qquad \frac{\partial^2 l}{\partial\alpha_0^T\partial\beta_1} = -\frac{n\delta_1^2\sigma^2}{B\sigma_\alpha^2\xi},$$

$$\frac{\partial^2 l}{\partial\alpha_1^T\partial\beta_1} = -\frac{\delta_1^2\sigma^2}{B\sigma_\alpha^2\xi}\sum_{i=1}^{n}S_{iq},$$

$$\begin{aligned}\frac{\partial^2 l}{\partial\alpha_2^T\partial\beta_1} &= \frac{\delta_1^2\sigma^2}{\xi}\sum_{i=1}^{n}\left(-\sum_{j=1}^{M+1}Y_{ij}Z_{ij} + \frac{\delta_1^2}{B\xi}\sum_{j=1}^{M+1}Z_{ij}\right) + \sum_{i=1}^{n}\sum_{j=1}^{M+1}(Y_{ij} - p_{ij})Z_{ij}\\ &\quad - \alpha_2^T\beta_1\sum_{i=1}^{n}\left\{\sum_{j=1}^{M+1}Z_{ij}Z_{ij}^T p_{ij} - \left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)\left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)^T\right\},\end{aligned}$$

$$\frac{\partial^2 l}{\partial\sigma_\alpha^2\partial\beta_1} = -\frac{\delta_1^3\sigma^2}{B^2\xi^2\sigma_\alpha^4}\sum_{i=1}^{n}\sum_{j=1}^{M+1}L_{ij},$$

$$\frac{\partial^2 l}{\partial\sigma^2\partial\beta_1} = \frac{\delta_1\sigma_W^2}{\xi^2}\sum_{i=1}^{n}\sum_{j=1}^{M+1}L_{ij}\left(Y_{ij} - \frac{\delta_1^2}{B\xi}\right) + \frac{\delta_1\sigma^2}{\xi}\left(-n\delta_1\beta_1 + \frac{n\delta_1^3\beta_1}{B\xi} + \frac{\delta_1^4}{B^2\xi^2\sigma_\alpha^2}\sum_{i=1}^{n}\sum_{j=1}^{M+1}L_{ij}\right),$$

$$\frac{\partial^2 l}{\partial\beta_2\partial\beta_2^T} = -\sum_{i=1}^{n}\left\{\sum_{j=1}^{M+1}Z_{ij}Z_{ij}^T p_{ij} - \left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)\left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)^T\right\},$$

$$\frac{\partial^2 l}{\partial\alpha_0\partial\beta_2^T} = -\beta_1\sum_{i=1}^{n}\left\{\sum_{j=1}^{M+1}Z_{ij}Z_{ij}^T p_{ij} - \left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)\left(\sum_{j=1}^{M+1}Z_{ij}p_{ij}\right)^T\right\},$$

$$\frac{\partial^2 l}{\partial\alpha_0^2} = -\zeta\delta_1 n(M+1), \qquad \frac{\partial^2 l}{\partial\alpha_1^T\partial\alpha_0} = -\zeta\delta_1(M+1)\sum_{i=1}^{n}S_{qi}, \qquad \frac{\partial^2 l}{\partial\alpha_2^T\partial\alpha_0} = -\zeta\delta_1\sum_{i=1}^{n}\sum_{j=1}^{M+1}Z_{ij},$$

$$\frac{\partial^2 l}{\partial\sigma_\alpha^2\partial\alpha_0} = -\frac{\delta_1^3}{B^2\xi^2\sigma_\alpha^2}\sum_{i=1}^{n}\sum_{j=1}^{M+1}L_{ij}, \qquad \frac{\partial^2 l}{\partial\sigma^2\partial\alpha_0} = \left(\frac{\delta_1^5}{\xi^3B^2\sigma_\alpha^2} - \frac{\delta_1^2\zeta}{\xi}\right)\sum_{i=1}^{n}\sum_{j=1}^{M+1}L_{ij} - n\beta_1\delta_1\zeta,$$

$$\frac{\partial^2 l}{\partial \alpha_1 \partial \alpha_1^T} = -\zeta\delta_1(M+1)\sum_{i=1}^{n} S_{qi}S_{qi}^T, \qquad \frac{\partial^2 l}{\partial \alpha_2^T \alpha_1} = -\zeta\delta_1 \sum_{i=1}^{n}\left(S_{qi}\sum_{j=1}^{M+1} Z_{ij}\right),$$

$$\frac{\partial^2 l}{\partial \sigma_\alpha^2 \partial \alpha_1} = -\frac{\delta_1^3}{B^2\xi^2\sigma_\alpha^4}\sum_{i=1}^{n}\left(S_{qi}\sum_{j=1}^{M+1} L_{ij}\right),$$

$$\frac{\partial^2 l}{\partial \sigma^2 \partial \alpha_1} = \left[\frac{\delta_1^5}{\xi^3 B^2\sigma_\alpha^2} - \frac{\delta_1^2\zeta}{\xi}\right]\sum_{i=1}^{n}\left(S_{qi}\sum_{j=1}^{M+1} L_{ij}\right) - \beta_1\delta_1\zeta\sum_{i=1}^{n} S_{qi},$$

$$\frac{\partial^2 l}{\partial \alpha_2 \partial \alpha_2^T} = -\zeta\delta_1\sum_{i=1}^{n}\sum_{j=1}^{M+1} Z_{ij}Z_{ij}^T - \beta_1^2\sum_{i=1}^{n}\left\{\sum_{j=1}^{M+1} Z_{ij}Z_{ij}^T p_{ij} - \left(\sum_{j=1}^{M+1} Z_{ij}p_{ij}\right)\left(\sum_{j=1}^{M+1} Z_{ij}p_{ij}\right)^T\right\},$$

$$\frac{\partial^2 l}{\partial \sigma_\alpha^2 \partial \alpha_2} = -\frac{\delta_1^3}{B^2\xi^2\sigma_\alpha^4}\sum_{i=1}^{n}\sum_{j=1}^{M+1} L_{ij}Z_{ij},$$

$$\frac{\partial^2 l}{\partial \sigma^2 \partial \alpha_2} = \left(\frac{\delta_1^5}{\xi^3 B^2\sigma_\alpha^2} - \frac{\delta_1^2\zeta}{\xi}\right)\sum_{i=1}^{n}\sum_{j=1}^{M+1} L_{ij}Z_{ij} - \beta_1\delta_1\zeta\sum_{i=1}^{n}\sum_{j=1}^{M+1} Y_{ij}Z_{ij},$$

$$\frac{\partial^2 l}{\partial \sigma_\alpha^2 \partial \sigma_\alpha^2} = \frac{n}{2\sigma_\alpha^4} + \frac{n}{2B^2\sigma_\alpha^8} - \frac{n}{B\sigma_\alpha^6} + \frac{\delta_1^2}{B^2\sigma_\alpha^6\xi^2}\left(\frac{1}{B\sigma_\alpha^2} - 1\right)\sum_{i=1}^{n}\left(\sum_{j=1}^{M+1} L_{ij}\right)^2,$$

$$\frac{\partial^2 l}{\partial \sigma^2 \partial \sigma_\alpha^2} = \frac{n(M+1)\delta_1^4}{2\sigma_\alpha^4 B^2\xi^2} - \frac{\delta_1^4}{B^3\xi^3\sigma_\alpha^6}\sum_{i=1}^{n}\left(\sum_{j=1}^{M+1} L_{ij}\right)^2 - \frac{\beta_1\delta_1^3}{B^2\sigma_\alpha^4\xi^2}\sum_{i=1}^{n}\sum_{j=1}^{M+1} L_{ij},$$

$$\begin{aligned}\frac{\partial^2 l}{\partial \sigma^2 \partial \sigma^2} = {} & -\frac{\delta_1^4}{\xi^3}\sum_{i=1}^{n}\sum_{j=1}^{M+1} L_{ij}^2 - \frac{2\delta_1^3\beta_1}{\xi^2}\sum_{i=1}^{n}\sum_{j=1}^{M+1} L_{ij}Y_{ij} + \left\{\frac{4\delta_1^5\beta_1}{B\xi^3} - \frac{2(M+1)\delta_1^7\beta_1}{B^2\xi^4}\right\}\sum_{i=1}^{n}\sum_{j=1}^{M+1} L_{ij} \\ & + \frac{\delta_1^6}{B\xi^4}\left\{\frac{\delta_1^4(M+1)^2}{B^2\xi^2} - \frac{3\delta_1^2(M+1)}{B\xi} + 3\right\}\sum_{i=1}^{n}\left(\sum_{j=1}^{M+1} L_{ij}\right)^2 - n\beta_1^2\delta_1\zeta + \frac{nM(M+1)\delta^6}{2B\xi^3} \\ & + \frac{n(M+1)\delta_1^4}{2B^2\xi^2\sigma_\alpha^2}\left(\frac{1}{\sigma_\alpha^2} + \frac{M\delta_1^2}{\xi}\right).\end{aligned}$$

Appendix B. Proof of Result 1

A1. The density function of the observed data is measurable for each parameter $\theta = (\beta^T, \psi^T)^T$ in Θ, a compact subset of the Euclidean space.

A2. $\partial l(\theta)/\partial\theta$ is a measurable function of $V_i \equiv (Y_{ij}, Z_{ij}, W_{ij},\ j = 1, \ldots, (M+1))$ for each $\theta \in \Theta$, and continuously differentiable function of θ for each V_i in the sample space.

A3. $|\partial^2 l(\theta)/\partial\theta_k\partial\theta_l|$ and $|\partial l(\theta)/\partial\theta_k \partial l(\theta)/\partial\theta_l|$ are dominated by functions integrable with respect to the data density for all V_i and for all $\theta \in \Theta$.

A4. $\theta^* = (\beta^{*T}, \psi^{*T})^T$ is in the interior of Θ, and $A^*_{\beta\cdot\psi}$ is non-singular.

Under the assumption that Θ is compact and $Q(\beta)$ is continuous in β, the restricted parameter space $\Theta_0 = \{\theta \in \Theta \cap Q(\beta) = 0\}$ is also compact. Under the regularity assumptions, $\tilde\theta \to \theta^*$ almost surely where $\theta^* \in \Theta_0$.

For the sake of convenience we will denote $(\partial l_i/\partial\beta)$ evaluated at $\beta = \tilde\beta$ and $\psi = \tilde\psi$ by $(\partial l_i(\tilde\beta, \tilde\psi)/\partial\beta)$. Let $\tilde\beta$ and $\tilde\psi$ be the estimators under H_0. Then

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\tilde\beta, \tilde\psi)}{\partial\beta} + \left\{\frac{\partial Q(\tilde\beta)}{\partial\beta}\right\}^T \hat\lambda = 0, \tag{B.8}$$

$$Q(\tilde\beta) = 0, \tag{B.9}$$

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\tilde\beta, \tilde\psi)}{\partial\psi} = 0, \tag{B.10}$$

where $\hat{\lambda}$ is the estimator of an $r \times 1$ vector of Lagrangian multipliers. Under the regularity assumptions, the estimators are consistent. Using the mean value theorem and assumption A2, we can write

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\tilde{\beta},\tilde{\psi})}{\partial\beta} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\beta^*,\psi^*)}{\partial\beta} + \bar{A}_{n\beta\beta}(\tilde{\beta}-\beta^*) + \bar{A}_{n\beta\psi}(\tilde{\psi}-\psi^*), \tag{B.11}$$

$$0 = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\tilde{\beta},\tilde{\psi})}{\partial\psi} = \frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\beta^*,\psi^*)}{\partial\psi} + \bar{A}_{n\psi\beta}(\tilde{\beta}-\beta^*) + \bar{A}_{n\psi\psi}(\tilde{\psi}-\psi^*), \tag{B.12}$$

where $\bar{\beta}_n$ lies on the line segment between β^* and $\tilde{\beta}$, and $\bar{\psi}$ lies on the line segment between ψ^* and $\tilde{\psi}$, and $\bar{(\cdot)}$ signifies that $(\cdot)$ is evaluated for $\beta=\bar{\beta}$ and $\psi=\bar{\psi}$. Eq. (B.12) implies that

$$(\tilde{\psi}-\psi^*) = -\bar{A}^{-1}_{n\psi\psi}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\beta^*,\psi^*)}{\partial\psi} + \bar{A}_{n\psi\beta}(\tilde{\beta}-\beta^*)\right\}. \tag{B.13}$$

Also, under H_0, we can re-write (B.9) as $Q(\tilde{\beta}) = Q(\beta^*) + \left\{\partial Q(\beta_n^\dagger)/\partial\beta\right\}(\tilde{\beta}-\beta^*) = 0$, where $\beta_n^\dagger$ is on the line segment joining β^* and $\tilde{\beta}$. From the above expression we obtain $\{\partial Q(\beta_n^\dagger)/\partial\beta\}(\tilde{\beta}-\beta^*) = 0$ by setting $Q(\beta^*) = 0$. Now replacing (B.12) into (B.8) and then using the expression (B.13) for $(\tilde{\psi}-\psi^*)$ we obtain

$$D_n(\bar{\beta},\bar{\psi})\begin{bmatrix}\frac{1}{n}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\beta\\ \frac{1}{n}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\psi\end{bmatrix} + \bar{A}_{n\beta\cdot\psi}(\tilde{\beta}-\beta^*) + \left\{\frac{\partial Q(\tilde{\beta})}{\partial\beta}\right\}^T\hat{\lambda} = 0, \tag{B.14}$$

where $D_n(\bar{\beta},\bar{\psi}) \equiv [I - \bar{A}_{n\beta\psi}\bar{A}^{-1}_{n\psi\psi}]$. Since $\bar{\beta}\to\beta^*$, $\bar{\psi}\to\psi^*$ almost surely, $\bar{A}_{n\beta\psi} \xrightarrow{\text{a.s}} A^*_{\beta\psi}$, $\bar{A}_{n\psi\psi} \xrightarrow{\text{a.s}} A^*_{\psi\psi}$, and $\bar{A}_{n\beta\cdot\psi}\to A^*_{n\beta\cdot\psi}$. Thus, $\bar{A}_{n\beta\cdot\psi}$ is also non-singular almost surely for large n. Now, premultiplying (B.14) by $\sqrt{n}\{\partial Q(\bar{\beta})/\partial\beta\}\bar{A}^{-1}_{n\beta\cdot\psi}$ and setting $\sqrt{n}\{\partial Q(\bar{\beta})/\partial\beta\}(\tilde{\beta}-\beta^*) = 0$ we get

$$\frac{\partial Q(\bar{\beta})}{\partial\beta}\bar{A}^{-1}_{n\beta\cdot\psi}D_n(\bar{\beta},\bar{\psi})\begin{bmatrix}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\beta\\ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\psi\end{bmatrix} + \sqrt{n}\frac{\partial Q(\bar{\beta})}{\partial\beta}\bar{A}^{-1}_{n\beta\cdot\psi}\left\{\frac{\partial Q(\bar{\beta})}{\partial\beta}\right\}^T\hat{\lambda} = 0.$$

By using Slutsky's theorem we get

$$\sqrt{n}\hat{\lambda} \overset{d}{=} -\left[\frac{\partial Q(\beta^*)}{\partial\beta}A^{*-1}_{\beta\cdot\psi}\left\{\frac{\partial Q(\beta^*)}{\partial\beta}\right\}^T\right]^{-1}\frac{\partial Q(\beta^*)}{\partial\beta}A^{*-1}_{\beta\cdot\psi}D(\beta^*,\psi^*)\begin{bmatrix}\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\beta\\ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\psi\end{bmatrix}.$$

Next, using the multivariate central limit theorem we obtain that under H_0, $\sqrt{n}\hat{\lambda} \sim \text{Normal}(0, Q)$, where the variance–covariance matrix is

$$Q \equiv \left[\frac{\partial Q(\beta^*)}{\partial\beta}A^{*-1}_{\beta\cdot\psi}\left\{\frac{\partial Q(\beta^*)}{\partial\beta}\right\}^T\right]^{-1}\frac{\partial Q(\beta^*)}{\partial\beta}A^{*-1}_{\beta\cdot\psi}D(\beta^*,\psi^*)CD^T(\beta^*,\psi^*)$$
$$\times A^{*-1}_{\beta\cdot\psi}\left\{\frac{\partial Q(\beta^*)}{\partial\beta}\right\}^T\left[\frac{\partial Q(\beta^*)}{\partial\beta}A^{*-1}_{\beta\cdot\psi}\left\{\frac{\partial Q(\beta^*)}{\partial\beta}\right\}^T\right]^{-1},$$

$$C = \begin{bmatrix}E[\{\partial l_i(\beta^*,\psi^*)/\partial\beta\}\{\partial l_i(\beta^*,\psi^*)/\partial\beta\}] & E[\{\partial l_i(\beta^*,\psi^*)/\partial\beta\}\{\partial l_i(\beta^*,\psi^*)/\partial\psi^T\}]\\ E[\{\partial l_i(\beta^*,\psi^*)/\partial\psi\}\{\partial l_i(\beta^*,\psi^*)/\partial\beta^T\}] & E[\{\partial l_i(\beta^*,\psi^*)/\partial\psi\}\{\partial l_i(\beta^*,\psi^*)/\partial\psi^T\}]\end{bmatrix}.$$

Therefore, under H_0

$$\begin{aligned}\mathcal{T}_2 &= \sqrt{n}\hat{\lambda}^T Q^{-1}\sqrt{n}\hat{\lambda}\\ &= \begin{bmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\beta \\ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\psi \end{bmatrix}^T D^T(\beta^*,\psi^*)A^{*-1}_{\beta\cdot\psi}\left\{\frac{\partial Q(\beta^*)}{\partial\beta}\right\}^T\left[\frac{\partial Q(\beta^*)}{\partial\beta}A^{*-1}_{\beta\cdot\psi}D(\beta^*,\psi^*)\right.\\ &\quad\left.\times CD^T(\beta^*,\psi^*)A^{*-1}_{\beta\cdot\psi}\left\{\frac{\partial Q(\beta^*)}{\partial\beta}\right\}^T\right]^{-1}\frac{\partial Q(\beta^*)}{\partial\beta}A^{*-1}_{\beta\cdot\psi}D(\beta^*,\psi^*)\begin{bmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\beta \\ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*)/\partial\psi \end{bmatrix}\end{aligned}$$

asymptotically follows a χ^2_r, a chi-square distribution with r degrees of freedom. The result follows after replacing the unknown parameters by their consistent estimators.

Appendix C. Proof of Result 2

This proof is similar to the previous proof. Thus, we just outline the main steps. The estimators are obtained by solving

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\tilde{\beta},\tilde{\psi},\hat{\phi})}{\partial\beta}+\left\{\frac{\partial Q(\tilde{\beta})}{\partial\beta}\right\}^T\hat{\lambda}=0, \tag{C.15}$$

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\tilde{\beta},\tilde{\psi},\hat{\phi})}{\partial\psi}=0, \tag{C.16}$$

along with Eq. (B.9), where $\hat{\phi}$ is a $\sqrt{m}$-consistent estimator of ϕ obtained from the external calibration data. Therefore, write

$$\sqrt{m}(\hat{\phi}-\phi^*)=\frac{1}{\sqrt{m}}\sum_{l=1}^{m}\vartheta_i+o_p(1), \tag{C.17}$$

where ϑ_is are 3-dimensional measurable functions such that $E(\vartheta_i)=0$, components of $E(\vartheta_i\vartheta_i^T)$ are finite and ϑ_i and $\vartheta_{i'}$ are independent for $i\neq i'$. Now, $(1/n)\sum_{i=1}^{n}\partial l_i(\tilde{\beta},\tilde{\psi},\hat{\phi})/\partial\beta$ can be approximated as

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\tilde{\beta},\tilde{\psi},\hat{\phi})}{\partial\beta}=\frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\beta^*,\psi^*,\phi^*)}{\partial\beta}+\bar{A}_{n\beta\beta}(\tilde{\beta}-\beta^*)+\bar{A}_{n\beta\psi}(\tilde{\psi}-\psi^*)+\bar{A}_{n\beta\phi}(\hat{\phi}-\phi^*). \tag{C.18}$$

Also, from (C.16) and using the mean-value theorem we get

$$(\tilde{\psi}-\psi^*)=-\bar{A}^{-1}_{n\psi\psi}\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{\partial l_i(\beta^*,\psi^*,\phi^*)}{\partial\psi}+\bar{A}_{n\psi\beta}(\tilde{\beta}-\beta^*)+\bar{A}_{n\psi\phi}(\hat{\phi}-\phi^*)\right\}. \tag{C.19}$$

Now using (C.17)–(C.19) into (C.15) we obtain

$$D_n(\bar{\beta}_n,\bar{\psi}_n,\hat{\phi})\begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*,\phi^*)/\partial\beta \\ \frac{1}{n}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*,\phi^*)/\partial\psi \\ \frac{1}{m}\sum_{l=1}^{m}\vartheta_l \end{bmatrix}+\bar{A}_{n\beta\cdot\psi}(\tilde{\beta}-\beta^*)+\left\{\frac{\partial Q(\tilde{\beta})}{\partial\beta}\right\}^T\hat{\lambda}=0, \tag{C.20}$$

where $D_n(\bar{\beta},\bar{\psi},\bar{\phi})\equiv[I\ -\bar{A}_{n\beta\psi}\bar{A}^{-1}_{n\psi\psi}\ \bar{A}_{n\beta\phi}-\bar{A}_{n\beta\psi}\bar{A}^{-1}_{n\psi\psi}\bar{A}_{n\psi\phi}]$. Define $D^*=\lim_{n\to\infty}D_n(\beta^*,\psi^*,\phi^*)$. Premultiplying both sides by $n^{1/2}\{\partial Q(\bar{\beta})/\partial\beta\}\bar{A}^{-1}_{n\beta\cdot\psi}$ and setting $n^{1/2}\{\partial Q(\bar{\beta})/\partial\beta\}(\tilde{\beta}-\beta^*)=0$ we obtain

$$\frac{\partial Q(\bar{\beta})}{\partial\beta}\bar{A}^{-1}_{n\beta\cdot\psi}D_n(\bar{\beta},\bar{\psi},\bar{\phi})\begin{bmatrix} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*,\phi^*)/\partial\beta \\ \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\partial l_i(\beta^*,\psi^*,\phi^*)/\partial\psi \\ \sqrt{\frac{n}{m}}\frac{1}{\sqrt{m}}\sum_{l=1}^{m}\vartheta_l \end{bmatrix}+\sqrt{n}\frac{\partial Q(\bar{\beta})}{\partial\beta}\bar{A}^{-1}_{n\beta\cdot\psi}\left\{\frac{\partial Q(\tilde{\beta})}{\partial\beta}\right\}^T\hat{\lambda}=0$$

which implies

$$\sqrt{n}\hat{\lambda}_n \stackrel{d}{=} \left[\frac{\partial Q(\beta^*)}{\partial \beta} A^{*-1}_{\beta\cdot\psi} \left\{\frac{\partial Q(\beta^*)}{\partial \beta}\right\}^T\right]^{-1} \frac{\partial Q(\beta^*)}{\partial \beta} A^{*-1}_{\beta\cdot\psi} D^* \begin{bmatrix} \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \partial l_i(\beta^*, \psi^*, \phi^*)/\partial \beta \\ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \partial l_i(\beta^*, \psi^*, \phi^*)/\partial \psi \\ \sqrt{\rho} \frac{1}{\sqrt{m}} \sum_{l=1}^{m} \vartheta_l \end{bmatrix}.$$

Using the multivariate central limit theorem we obtain $\sqrt{n}\hat{\lambda} \sim \text{Normal}(0, Q)$, where

$$\begin{aligned} Q &= \text{var}(\sqrt{n}\hat{\lambda}) = \left[\frac{\partial Q(\beta^*)}{\partial \beta} A^{*-1}_{\beta\cdot\psi} \left\{\frac{\partial Q(\beta^*)}{\partial \beta}\right\}^T\right]^{-1} \frac{\partial Q(\beta^*)}{\partial \beta} A^{*-1}_{\beta\cdot\psi} D^* C D^* A^{*-1}_{\beta\cdot\psi} \left\{\frac{\partial Q(\beta^*)}{\partial \beta}\right\}^T \\ &\quad \times \left[\frac{\partial Q(\beta^*)}{\partial \beta} A^{*-1}_{\beta\cdot\psi} \left\{\frac{\partial Q(\beta^*)}{\partial \beta}\right\}^T\right]^{-1}, \quad \text{and} \\ C &= \begin{bmatrix} E[\{\partial l_i/\partial \beta\}\{\partial l_i/\partial \beta\}] & E[\{\partial l_i/\partial \beta\}\{\partial l_i/\partial \psi^T\}] & 0 \\ E[\{\partial l_i/\partial \psi\}\{\partial l_i/\partial \beta^T\}] & E[\{\partial l_i/\partial \psi\}\{\partial l_i/\partial \psi^T\}] & 0 \\ 0 & 0 & \rho E[\vartheta_l \vartheta_l^T] \end{bmatrix}_{\beta=\beta^*, \psi=\psi^*, \phi=\phi^*}. \end{aligned}$$

Hence Result 2 follows. In computing the test statistics we replace the expectations by the empirical averages, and replace β^*, ψ^*, and ϕ^* by the corresponding consistent estimators under H_0.

References

[1] B.G. Armstrong, A.S. Whittemore, G.R. Howe, Analysis of case-control data with covariate measurement error: application to diet and colon cancer, Statistics in Medicine 8 (1989) 1151–1163.
[2] E. Budtz-Jørgensen, N. Keiding, P. Grandjean, P. Weihe, R.F. White, Consequences of exposure measurement errors for confounder identification in environmental epidemiology, Statistics in Medicine 22 (2003) 3089–3100.
[3] R, J. Carroll, D. Ruppert, L.A. Stepanski, C. Crainiceanu, Measurement Error in Nonlinear Models: A Modern Perspective, second ed., Chapman and Hall, New York, 2006.
[4] M. de Castro, M. Galea, H. Bolfarine, Hypothesis testing in an errors-in-variables model with heteroscedastic measurement errors, Statistics in Medicine 27 (2008) 5217–5234.
[5] A. Guolo, A.R. Brazzale, A simulation-based comparison of techniques to correct for measurement error in matched case-control studies, Statistics in Medicine 27 (2008) 3755–3775.
[6] M. Haftenberger, T. Heuer, C. Heidemann, F. Kube, C. Krems, G. Mensink, Relative validation of a food frequency questionnaire for national health and nutrition monitoring, Nutrition Journal 9 (2010) 36.
[7] M. Jain, G.M. Cook, F.G. Davis, M.G. Grace, G.R. Howe, A.B. Miller, A case-control study on diet and colo-rectal cancer, International Journal of Cancer 26 (1980) 757–768.
[8] L.M. McShane, D.N. Midthune, J.F. Dorgan, L.S. Freedman, R.J. Carroll, Covariate measurement error adjustment for matched case-control studies, Biometrics 57 (2001) 62–73.
[9] H. Murad, L.S. Freedman, Estimating and testing interactions in linear regression models when explanatory variables are subject to classical measurement error, Statistics in Medicine 26 (2007) 4293–4310.
[10] R.L. Prentice, et al., Low-fat dietary pattern and risk of invasive breast cancer, Journal of the American Medical Association 295 (2006) 629–642.
[11] G.A. Satten, L.L. Kupper, Inferences about exposure-disease associations using probability-of-exposure information, Journal of the American Statistical Association 88 (1993) 200–208.
[12] S. Sinha, B. Mukherjee, M. Ghosh, B.K. Mallick, R.J. Carroll, Semiparametric Bayesian analysis of matched case-control studies with missing exposure, Journal of the American Statistical Association 100 (2005) 591–601.
[13] L.A. Stefanski, R.J. Carroll, Conditional scores and optimal scores in generalized linear measurement error models, Biometrika 74 (1987) 703–716.
[14] L.A. Stefanski, R.J. Carroll, Score tests in generalized linear measurement error models, Journal of the Royal Statistical Society, Series B 52 (1990) 345–359.
[15] T.D. Tosteson, A.A. Tsiatis, The asymptotic relative efficiency of score tests in a generalized linear model with surrogate covariates, Biometrika 75 (1988) 507–514.
[16] H. White, Maximum likelihood estimation of misspecified models, Econometrica 50 (1982) 1–24.