Stat Biosci (2014) 6:1–18
DOI 10.1007/s12561-012-9074-5

On Method of Moments Estimation in Linear Mixed Effects Models with Measurement Error on Covariates and Response with Application to a Longitudinal Study of Gene-Environment Interaction

Taraneh Abarin · He Li · Liqun Wang · Laurent Briollais

Received: 13 October 2011 / Accepted: 18 October 2012 / Published online: 13 November 2012

Abstract We study a linear mixed effects model for longitudinal data, where the response variable and covariates with fixed effects are subject to measurement error. We propose a method of moment estimation that does not require any assumption on the functional forms of the distributions of random effects and other random errors in the model. For a classical measurement error model we apply the instrumental variable approach to ensure identifiability of the parameters. Our methodology, without instrumental variables, can be applied to Berkson measurement errors. Using simulation studies, we investigate the finite sample performances of the estimators and show the impact of measurement error on the covariates and the response on the estimation procedure. The results show that our method performs quite satisfactory, especially for the fixed effects with measurement error (even under misspecification of measurement error model). This method is applied to a real data example of a large birth and child cohort study.

Keywords Measurement error · Mixed effects model · Longitudinal data · Instrumental variable · Method of moments · Gene-environment interaction

This study has been partially supported by NSERC.

Electronic supplementary material The online version of this article (doi:10.1007/s12561-012-9074-5) contains supplementary material, which is available to authorized users.

T. Abarin (✉)
Department of Mathematics and Statistics, Memorial University, St. John's, NL, Canada
e-mail: tabarin@mun.ca

H. Li · L. Wang
Department of Statistics, University of Manitoba, Winnipeg, Canada

L. Briollais
Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Canada

1 Introduction

It is widely recognized that both genetic and environmental factors influence almost every complex trait, and also that these genetic risks result from a large number of individually small effect sizes. Genetic factors are also likely to interact with environmental factors. Studies of gene-environment interactions (GEIs) can improve the accuracy and precision of the assessment of both genetic and environmental influences. There are many GEI studies using longitudinal data (e.g. [16, 26, 28, 34]) that assume that all the environmental factors are measured accurately. Other studies such as Burton et al. [13], Thomas [31], Marchand and Wilkens [24], and Wong et al. [39] only acknowledge the presence of measurement error (ME) in the covariates and its impact on GEI estimates, but they did not adjust for the bias that ME causes on the estimators.

For the aforementioned reasons, the maximum likelihood [21] is still the most common method of estimating GEI effects in longitudinal studies. This methodology ignores ME and has restrictions on the distributional assumptions on model error terms and random effects.

For independent observations, a comprehensive review of ME correction methods is given in Fuller [17] and Carroll et al. [14]. In an early study, Tosteson et al. [32] used repeated measures on the covariates X and a normality assumption on ME to correct the bias of the estimators. There are several methodologies suggested to deal with ME in linear and nonlinear models. Some authors study regression calibration and simulation extrapolation [12, 37, 38] for generalized linear mixed effects models with ME. These two methods are only approximately consistent. A similar problem arises in the nonparametric method of Carroll and Wand [15]. A major difficulty in applying these approaches is that the unobserved covariate has repeated measures which are likely correlated. Likelihood-based methods have also been investigated by Buonaccorsi [8] and Buonaccorsi et al. [12]. Generally, likelihood approaches suffer from restrictive distributional assumptions on the random effects, ME, covariates with ME and the model error term. Since random effects, error-prone covariates and ME are unobservable, likelihood-based approaches might not be realistic. They are also known to be sensitive to distributional assumptions and outliers, even in mixed effects models without ME. Non- or semi-parametric approaches have also been considered, for example, by Tsiatis and Davidian [33], who relax the normality assumption of the random effects but not of the ME. Pan et al. [25] also considered an approach that relaxes the distributional assumption on the underlying error-prone covariates but not on the ME nor on the response.

ME on response has been considered in some studies [6, 18, 19]. Carroll et al. [14] considered an unbiased response while others considered biased response and corrected for the bias [9, 10].

Motivated by the Raine study, a large birth and child cohort in Australia, we propose a method of moments estimation for a linear mixed effects model (LMEM) for longitudinal data with ME on covariates and response based on the conditional moments of the response given the observed covariates. The proposed approach does not require any parametric assumptions on the distributions of the covariates with ME, random effects, and measurement errors, which are difficult to check in practice. Wang [35] applied a methodology to a nonlinear homoscedastic regression model with Berkson type measurement error only on the covariates and showed that it is identifiable without extra information. Li [22] applied a similar method called second-order least squares to the LMEM and to the nonlinear mixed effects model without ME. In addition, Li and Wang [23] investigated various computational issues related to the proposed estimator, such as redescending property and its robustness against data contamination.

Classical ME models usually need extra information such as replicate data or instrumental variables in order to be identifiable. Abarin and Wang [2] suggested a semi-parametric method for estimating parameters of generalized linear regression models with the classical ME model. They considered a model for instrumental variables to deal with the identifiability issue. Following the ideas proposed by Abarin and Wang [2], and Li and Wang [23], we propose an instrumental variable model for a linear mixed effects model with ME on covariates and response. Here we show that, using moment equations, we can estimate all the parameters of interest in the model. Using simulation studies, we investigate the finite sample performances of the estimators and show the impact of ME on the covariates and response on the estimation procedure. We also apply this method to a real example from a large birth and child cohort study in Australia (Raine).

This paper is organized as follows. In Sect. 2, we introduce the LMEM for a response with ME and also classical ME models for the error-prone covariates. We also present the method of moments estimators and show how instrumental variables can be applied to the estimation procedure. In Sect. 3, we examine an alternative model with a Berkson ME on covariates. We investigate the finite sample performances of the proposed estimators in comparison with the naive MLE both for the Berkson and for the classical ME models in Sect. 4. In this section, we also examine how the method of moments estimators behaves when we have misspecification in the ME model on covariates. The estimation approaches are illustrated in Sect. 5 with the analysis of a data set from the Raine study. Finally, a summary and discussion are provided in Sect. 6.

2 Linear Mixed Effects Model with Measurement Error

Motivated by the Raine study, we define a linear mixed measurement error model for the jth observation on the ith individual as

$$Y_{ij} = A_{ij}\beta + B_{ij}\nu_i + \delta_{ij}, \quad i = 1, \dots, n, \ j = 1, \dots, N_i, \tag{1}$$

where Y_{ij} is the jth response for the ith observation, $A_{ij} = (1, X'_{ij}, Z'_{ij})$ is the design matrix for fixed effects, $B_{ij} = (1, Z'_{ij})$ is the design matrix for random effects, $X_{ij} \in \mathbb{R}^{p_x}$ is the vector of predictors with measurement error, $Z_{ij} \in \mathbb{R}^{p_z}$ is the vector of accurately measured covariates, $\beta = (\beta_0, \beta'_x, \beta'_z)'$ is the vector of fixed effects, and δ_{ij}s are mutually independent error terms (see Summary and Discussion for further details regarding this assumption) with zero mean and equal variances σ^2_δ. Here we used Z for both fixed effects and random effects without ME. Obviously, in practice, these two sets of variables might be different, as in the example in Sect. 5. This was done to simplify the notations.

In model (1), $\nu_i = (\nu_{0i}, \nu'_{zi})'$ are vectors of random effects, independent of observed covariates, associated with an individual i, and have a general distribution with mean 0 and covariance matrix Σ_ν. Here, we could include an interaction term between X and one or more variables in the design matrix B. It does not make any difference in our methodology, so for simplicity, we work with model (1). Further, suppose that X_{ij} is unobservable. Instead, we observe W_{ij} defined as

$$W_{ij} = X_{ij} + \varepsilon_{ij}, \tag{2}$$

where ε is a random measurement error with mean 0 (see Summary and Discussion for further details regarding this assumption) and independent from X. This is called a classical additive ME model [14], which is the most common model for ME on covariates. This model allows random effects on the covariates that have no measurement errors. We also suggest a classical measurement error for the response as

$$S_{ij} = Y_{ij} + \xi_{ij}, \tag{3}$$

where ξ is a random measurement error with mean 0 and covariance matrix $\sigma_\xi^2 I$. This model also assumes that ξ is independent from Y. Moreover, we assume that W is surrogate, meaning that given the true covariates, W does not provide any extra information regarding the distribution of the response. As we mentioned earlier, this model selection is motivated by our study on the Raine study. More details regarding this model selection is provided in Sect. 5.

It is not obvious how to determine the parameters for which the naive estimators are inconsistent unless we have more assumptions on the model. For example, Carroll et al. [14] showed that if X is normally distributed, if a classical additive error model holds, and if X and Z are independent, then the naive estimator will be consistent only for the fixed and the random effects corresponding to Z. Unlike the impact of ME on X, it is straightforward to see how ME affects the response under the classical additive ME model in (3). Since this model assumes that Y and ξ are independent and ξ has mean 0, the naive estimator that uses S instead of Y remains unbiased. However, ignoring ME on Y and simply assuming that the error gets absorbed into the model error is a myth. Even an unbiased ME on response increases the variability of the fitted model [14].

Following Schennach [30], Abarin and Wang [2], and Wang and Hsiao [36], we assume that an instrumental variable $V \in \mathbb{R}^q$ is available and related to X through

$$X_{ij} = GV_{ij} + U_{ij}, \tag{4}$$

where G is a row full rank matrix of unknown parameters and U is independent from V and ε, with mean 0 and variance–covariance matrix $\sigma_u^2 I$. Further, all random errors and random effects $(\delta, \epsilon, \xi, U, \nu)$ are mutually independent and have conditional mean 0 given (V, Z). Since there is no assumption concerning the functional forms of the distributions of X, δ, ε, and ξ, the model (1)–(3) is semi-parametric. Substituting (4) into (2) results in the linear regression equation

$$E(W_{ij}|V) = GV_{ij}. \tag{5}$$

In practice, one can estimate G consistently, either using an external independent sample or a subset of the main sample, and estimate the other parameters using the rest of the sample. We let $X_i = (X'_{i1}, X'_{i2}, \dots, X'_{iN_i})$, and W_i, V_i analogously, so G can be estimated by

$$\hat{G} = \left(\sum_{i=1}^{n} W_i V_i'\right)\left(\sum_{i=1}^{n} V_i V_i'\right)^{-1}. \tag{6}$$

Based on Abarin and Wang [2] and Li and Wang [23], we can write three sets of conditional moments as

$$E(S_{ij}|V, Z) = D_{ij}\beta, \tag{7}$$

$$\begin{aligned} E(S_{ij}S_{ik}|V, Z) = {} & E(S_{ij}|V, Z)E(S_{ik}|V, Z) + B_{ij}\Sigma_\nu B'_{ik} \\ & + \sigma_{u_{ijk}}\beta'_x\beta_x + \sigma_{\delta_{ijk}} + \sigma_{\xi_{ijk}}, \end{aligned} \tag{8}$$

and

$$E(S_{ij}W_{ik}|V, Z) = D_{ij}\beta G V_{ik} + \sigma_{u_{ijk}}\beta_x, \tag{9}$$

where $D_{ij} = (1, V'_{ij}G', Z'_{ij})$, and $\sigma_{u_{ijk}} = \sigma_u^2$ if $j = k$ and 0 otherwise. Similarly, $\sigma_{\delta_{ijk}} = \sigma_\delta^2$ and $\sigma_{\xi_{ijk}} = \sigma_\xi^2$ if $j = k$ and 0 otherwise. Schennach [30], Wang and Hsiao [36] have shown that for a general model with independent cross-sectional data and ME only on the covariates, the parameters can be identified using V and these moment equations, provided certain regularity conditions hold. As we can see in the second conditional moment, $\sigma_{\delta_{ijk}}$ and $\sigma_{\xi_{ijk}}$ appear together and therefore are not identifiable. In practice, we can estimate $\sigma_{\xi_{ijk}}$ using an external independent sample or a validation subsample. Although in practice σ_ξ^2 and σ_u^2 are not usually the parameters of interest, estimating σ_ξ^2 guarantees a consistent estimator for σ_δ^2. In this case, after estimating σ_ξ^2, σ_δ^2 can be estimated by subtracting $\hat{\sigma}_\xi^2$ from the total variability of δ and ξ.

In this model, the observed vector of variables is (S, Z, W, V). Our interest is to estimate $\gamma = (\beta', \mathrm{vec}(\Sigma_\nu)', \sigma_\delta^2, \sigma_u^2)'$.

Define $\rho_i(\gamma)$ to be $(S_{ij} - E(S_{ij}|V, Z), 1 \le j \le N_i, S_{ij}S_{ik} - E(S_{ij}S_{ik}|V, Z), S_{ij}W_{ij} - E(S_{ij}W_{ij} \mid V, Z), 1 \le j \le k \le N_i)'$. Then the method of moments estimator (MME) for γ is defined as

$$\hat{\gamma}_{\mathrm{MME}} = \arg\min_{\gamma\in\Gamma} Q_n(\gamma) = \arg\min_{\gamma\in\Gamma} \sum_{i=1}^{n} \rho'_i(\gamma) H_i \rho_i(\gamma), \tag{10}$$

where H_i is a nonnegative definite matrix which may depend on V and Z.

We should also mention that adding interaction terms between X and one or more variables in the design matrix B does not affect our estimation procedure. Wang et al. [32] showed that the naive estimator of the coefficients corresponding to the variable with ME for Gaussian data and a classic ME model are asymptotically biased. Since the naive estimator of the coefficients corresponding to the variable with ME is biased, intuitively we can assume that the estimator of the interaction term with it is also biased.

Theorem 2.1 *Under some regularity conditions,* $\hat{\gamma}_{\mathrm{MME}}$ *is strongly consistent and* $\sqrt{n}(\hat{\gamma}_{\mathrm{MME}}-\gamma_0)\xrightarrow{L} N(0,\kappa^{-1}C\tau C'\kappa^{-1})$ *as* $n\to\infty$, *where* γ_0 *is the true value for* γ. *Here,*

$$\begin{aligned}
C &= \left[I, E\left(\frac{\partial\rho'(\gamma_0)}{\partial\gamma}H\frac{\partial\rho(\gamma_0)}{\partial\psi'}\right)\left(EVV'\otimes I\right)^{-1}\right],\\
\tau &= \begin{pmatrix}\tau_{11} & \tau_{12}\\ \tau_{12}' & \tau_{22}\end{pmatrix},\\
\tau_{11} &= E\left[\frac{\partial\rho'(\gamma_0)}{\partial\gamma}H\rho(\gamma_0)\rho'(\gamma_0)H\frac{\partial\rho(\gamma_0)}{\partial\gamma'}\right],\\
\tau_{12} &= E\left[\frac{\partial\rho'(\gamma_0)}{\partial\gamma}H\rho(\gamma_0)\left((W-GV)'\otimes V'\right)\right],\\
\tau_{22} &= E\left[VV'\otimes(W-GV)(W-GV)'\right],
\end{aligned}$$

and

$$\kappa = E\left[\frac{\partial\rho'(\gamma_0)}{\partial\gamma}H\frac{\partial\rho(\gamma_0)}{\partial\gamma'}\right]. \tag{11}$$

Proof Proofs can be found in the Supplementary Materials. □

We notice that estimating G creates extra variation in the final estimator. If G was known, τ would be reduced to

$$\tau = E\left[\frac{\partial\rho'(\gamma_0)}{\partial\gamma}H\rho(\gamma_0)\rho'(\gamma_0)H\frac{\partial\rho(\gamma_0)}{\partial\gamma'}\right]. \tag{12}$$

The theorem actually shows that MME gets closer to the true value of the parameter, when the sample size increases. Therefore, the finite sample bias of this method decreases with an increase in the sample size. This is not the case with the naive estimator, however, which does not decrease with an increase in the sample size.

The above asymptotic covariance matrix depends on the weighting matrix H_i. It is of interest to choose an appropriate matrix H_i to obtain the most efficient estimator. It can be shown [4] that the most efficient choice of weight is $H_i = F_i^{-1}$, where $F_i = E(\rho_i\rho_i'|V,Z)$, which leads to the covariance matrix

$$E\left[\frac{\partial\rho_i'(\gamma_0)}{\partial\gamma}F_i^{-1}\frac{\partial\rho_i(\gamma_0)}{\partial\gamma'}\right]^{-1}. \tag{13}$$

In practice, F_i is a function of unknown parameters and must be estimated. This can be done using the following procedure. First, minimize $Q_n(\gamma)$ with the identity matrix to obtain the first-stage estimator $\hat{\gamma}^{(1)}_{\mathrm{MME}}$. Then, estimate F_i by $\hat{F}_i = \frac{1}{n}\sum_{i=1}^n\rho_i(\hat{\gamma}^{(1)}_{\mathrm{MME}})\rho_i'(\hat{\gamma}^{(1)}_{\mathrm{MME}})$ or, alternatively, by a nonparametric estimator and then minimize $Q_n(\gamma)$ again with $H_i = \hat{F}_i^{-1}$ to obtain the second-stage estimator $\hat{\gamma}^{(2)}_{\mathrm{MME}}$.

Since $\hat{F}_i$ is consistent for F_i, the asymptotic covariance of $\hat{\gamma}^{(2)}_{\text{MME}}$ is given by (13). Consequently $\hat{\gamma}^{(2)}_{\text{MME}}$ is asymptotically more efficient than the first-stage estimator $\hat{\gamma}^{(1)}_{\text{MME}}$. The iteration process continues until convergence. As can be seen, F_i is the same for all the individuals. While there might be other ways of computing the weighting matrix, ours is probably the most convenient one. There are other methods to compute the weighting matrices for each individual. In practice, matrix F_i might be singular or near singular. To ease computational difficulties, it is more practical to use a diagonal form of (13). In theory, this matrix is not as efficient as the optimum weighting matrix; however, it has been shown in the numerical studies by Li and Wang [23] that the efficiency loss is considerably small.

3 Berkson Measurement Error Models for Covariates

A Berkson ME model [5] for X_{ij} is defined as

$$X_{ij} = W_{ij} + \omega_{ij}, \tag{14}$$

where ω is a random measurement error with mean 0 and variance–covariance matrix $\sigma^2_\omega I$ independent from W. Here we assume that the true covariate has more variability than the observed covariate. Although (14) might look similar to (2), they are actually very different. "Basically, if the covariate with ME is necessarily measured uniquely to an individual, and especially if that measurement can be replicated, the ME model is classical. On the other hand, if all individuals in a small group are given the same value of the covariate with ME, but the true covariate is particular to an individual, then the ME model is Berkson" [14]. Moreover, in (14), W and ω are independent, which results in $\text{var}(X) > \text{var}(W)$, while for a classical ME model, $\text{var}(W) > \text{var}(X)$.

Substituting (3) and (14) into (1), we have

$$S_{ij} = C_{ij}\beta + B_{ij}\nu_i + \delta_{ij} + \omega'_{ij}\beta_x + \xi_{ij}, \quad i = 1, \dots, n,\ j = 1, \dots, N_i, \tag{15}$$

where $C_{ij} = (1, W'_{ij}, Z'_{ij})$. Comparing the parameters in (1) to those in (15), we can see that the naive estimators (which ignore ME) of fixed effects and the variance components of Σ_ν are consistent. The only variance component for which the naive estimator is inconsistent is σ^2_δ. Since the error in (15) is $\delta_{ij} + \omega'_{ij}\beta_x + \xi_{ij}$, the naive estimator is consistent for $\sigma^2_\delta + \sigma^2_\omega\beta'_x\beta_x + \sigma^2_\xi$ instead of σ^2_δ. Even if we estimate σ^2_ξ using either an external sample or a subset of the main sample in advance, σ^2_δ is still not identifiable. This estimator is crucial for predicting the response using the true covariates. More specifically, in testing hypotheses, the presence of ME on some of the covariates overestimates σ^2_δ potentially causing "false negative" results. Similarly, all random errors and random effects $(\delta, \omega, \xi, \nu)$ are mutually independent and have conditional mean 0 given (X, Z). To illustrate how an interaction term between X and Z can affect the estimation procedure, for now, we assume that our model is

$$\begin{aligned} Y_{ij} &= \beta_0 + \beta_x X_{ij} + \beta_z Z_{ij} + \beta_{zx} X_{ij} Z_{ij} + B_{ij}\nu_i + \delta_{ij}, \\ &\quad i = 1, \dots, n,\ j = 1, \dots, N_i. \end{aligned} \tag{16}$$

Substituting (3) and (14) into (16), we have

$$S_{ij} = \beta_0 + \beta_x W_{ij} + \beta_z Z_{ij} + \beta_{zx} W_{ij} Z_{ij} + B_{ij}\nu_i \\ + \delta_{ij} + \beta_x \omega_{ij} + \beta_{zx}\omega_{ij} Z_{ij} + \xi_{ij}, \quad i = 1, \dots, n, \; j = 1, \dots, N_i.$$

This shows that the naive estimator of the effect on the interaction term is unbiased and that adding an interaction term in the model only increases the variability of δ. This also shows that the interaction term does not affect MME either since we treat the interaction term as a new effect.

Now we show that using only the first two moment equations, we can estimate all the parameters of interest. Following the methodology of Wang [35] under the model assumptions, the conditional mean of the observed response S_{ij} given observed variables is given by

$$E(S_{ij}|W, Z) = C_{ij}\beta, \tag{17}$$

and the moments of S_{ij} given W and Z are

$$E(S_{ij}S_{ik}|W, Z) = E(S_{ij}|W, Z)E(S_{ik}|W, Z) + B_{ij}\Sigma_\nu B'_{ik} \\ + \sigma_{\omega_{ijk}}\beta'_x\beta_x + \sigma_{\delta_{ijk}} + \sigma_{\xi_{ijk}}, \tag{18}$$

where $\sigma_{\omega_{ijk}} = \sigma^2_\omega$ if $j = k$, and 0 otherwise. For this case, we define $\rho_i(\gamma)$ to be $(S_{ij} - E(S_{ij}|W, Z), 1 \le j \le N_i, S_{ij}S_{ik} - E(S_{ij}S_{ik}|W, Z), 1 \le j \le k \le N_i)'$, and the method of moments estimator (MME) for γ is defined as

$$\hat{\gamma}_{\mathrm{MME}} = \arg\min_{\gamma\in\Gamma} Q_n(\gamma) = \arg\min_{\gamma\in\Gamma} \sum_{i=1}^{n} \rho'_i(\gamma)H_i\rho_i(\gamma), \tag{19}$$

where H_i is a nonnegative definite matrix which may depend on W and Z. Here again, we need to have additional information on $\sigma_{\xi_{ijk}}$ in order to have a consistent estimate for $\sigma_{\delta_{ijk}}$. Similar to the classic model for X, it can be shown that $\hat{\gamma}_{\mathrm{MME}}$ is strongly consistent and asymptotically normally distributed with mean 0 and the covariance matrix given by $A^{-1}BA^{-1}$ as $n \to \infty$, where γ_0 is the true value for γ. Here,

$$B = E\left[\frac{\partial\rho'(\gamma_0)}{\partial\gamma} H\rho(\gamma_0)\rho'(\gamma_0)H\frac{\partial\rho(\gamma_0)}{\partial\gamma'}\right],$$

and

$$A = E\left[\frac{\partial\rho'(\gamma_0)}{\partial\gamma} H\frac{\partial\rho(\gamma_0)}{\partial\gamma'}\right]. \tag{20}$$

The proofs of these two properties of MME can be found in the Supplementary Materials.

4 Simulation Studies

In this section, we carry out simulation studies with different scenarios to show the impact of ME on covariates only, or on both covariates and the response, using the method of moments and naive maximum likelihood estimators. We are also interested in examining the impact of the sample size on the estimators and their finite sample behavior. We examined these issues under both classical and Berkson ME models. Moreover, we investigated the sensitivity of MME under misspecification of the ME model.

4.1 The Set Up

We considered the following simple LMEM with two different sample sizes $n = 100$ and $n = 300$:

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + \nu_{0i} + \delta_{ij}, \quad j = 1, \ldots, 4. \tag{21}$$

The random intercept ν_{0i} was generated from a normal distribution with mean 0 and variance 0.25, and δ_{ij} was generated from a standard normal distribution. The parameters of interest are $\beta_0 = 8$, $\beta_1 = 2$, $\sigma^2_{\nu_0} = 0.25$, and $\sigma^2_\delta = 1$.

4.2 Case One—Classical ME

The instrumental variable, U, and ε, were all generated from a standard normal distribution. Then we generated the unobserved X through an instrumental model that describes the relationship between X and V according to

$$X_i = 1.2 + 0.4 \cdot V_i + U_i, \tag{22}$$

where U is independent from V and ε. The values of 1.2 and 0.4 are only arbitrary and assumed to be known. We could therefore generate W according to the following calorical ME model:

$$W_i = X_i + \varepsilon_i.$$

Here, we assume one observation for every individual.

4.3 Case Two—Berkson ME

We generated the unobserved X according to

$$X_i = W_i + \omega_i,$$

where W and ω are independent and have standard normal distributions.

For both cases, error-prone response was generated according to model (3), where ξ_{ij} has a standard normal distribution. In order to be able to estimate σ^2_δ when we have ME on response, we treated σ^2_ξ as known.

The reason for generating all the variables from normal distributions is to make a fair comparison with the naive MLE. We could examine other distributions because

MME works well for them. However, since there are no standard functions available for the naive MLE for those distributions, we avoided them. In fact, MME shows superiority to MLE in a misspecified censored model without ME [3]. So, we would expect to see that it performs better than MLE in this model as well.

4.4 Case Three—Misspecified ME

To examine the sensitivity of the MME when we have misspecification in the ME model on covariates, we assumed that the ME model for X is classical when it was actually Berkson. A classical ME is the most frequently used model, so most likely to be chosen by default when one does not know the details of the design of a study. In order to ensure that all the relationships between the variables are satisfied, we generated U and ω independently from a standard normal distribution and then generated W and V from a bivariate normal distribution with mean vector $(0.2, 0)'$ and variances of 3.96 and 1.4, respectively. The covariance between the two variables can be easily calculated based on (22) and the classical ME for X. In the last step, we generated X from a Berkson model. All values are only arbitrary.

For each of the sample sizes, $R = 1000$ Monte Carlo replicates were simulated and the Monte-Carlo mean estimates and root mean squared errors (RMSE) of the estimators were computed. All computations were done in R, and ML (naive) estimates were obtained from the `lmer` package. The MME was computed using fully estimated optimal weight. To determine how well the methods perform, we present the bias and RMSE of the estimators. To eliminate some potential nonlinear numerical optimization problems in the determination of the starting points, the true parameter values were used as starting values for the minimization and the optimal weight calculation for the MME method. For faster results, it is better to have a consistent estimator as a starting point for MME. In practice, although it is not consistent, the naive estimator can be used.

4.5 Results

Tables 1–5 summarize the results of the simulations.

4.5.1 Case One—Classical ME

Table 1 shows the results for a classical model when we have ME on X only. Table 2 shows the results for a classical model when we have ME on both X and Y. As

Table 1 Bias (RMSE) of the MLE and MME estimators based on 1000 Monte-Carlo simulations for the classical ME model with ME on X

True value	$n = 100$		$n = 300$		CP
	MLE	MME	MLE	MME	
$\beta_0 = 8$	0.1031(0.1499)	0.0060(0.0921)	0.1003(0.1197)	−0.0001(0.0631)	94.8 %
$\beta_1 = 2$	−0.5068(0.5093)	−0.0406(0.0669)	−0.5050(0.5059)	−0.0176(0.0436)	95.0 %
$\sigma^2_{\nu_0} = 0.25$	0.0078(0.1786)	0.0079(0.3487)	−0.0021(0.1113)	0.0063(0.2250)	56.8 %
$\sigma^2_{\delta} = 1$	2.9808(2.9963)	0.0121(0.4393)	2.9961(3.0019)	0.0240(0.3018)	37.9 %

Table 2 Bias (RMSE) of the MLE and MME estimators based on 1000 Monte-Carlo simulations for the classical ME model with ME on both X and Y

True value	$n=100$		$n=300$		CP
	MLE	MME	MLE	MME	
$\beta_0=8$	0.0952(0.1534)	0.0024(0.1036)	0.0996(0.1214)	−0.0001(0.0690)	93.9 %
$\beta_1=2$	−0.5035(0.5066)	−0.0441(0.0759)	−0.5044(0.5055)	−0.0213(0.0511)	94.9 %
$\sigma^2_{\nu_0}=0.25$	0.0221(0.2166)	0.0352(0.4088)	0.0003(0.1323)	0.0194(0.2536)	45.6 %
$\sigma^2_\delta=1$	3.9773(3.9960)	−0.0605(0.5110)	3.9874(3.9943)	0.0031(0.3433)	39.6 %

Table 3 Bias (RMSE) of the MLE and MME estimators based on 1000 Monte-Carlo simulations for the Berkson ME model with ME on X

True value	$n=100$		$n=300$		CP
	MLE	MME	MLE	MME	
$\beta_0=8$	0.0021(0.1587)	0.0009(0.1526)	−0.0006(0.0904)	−0.0062(0.0952)	95.1 %
$\beta_1=2$	0.0004(0.0801)	−0.0322(0.0782)	−0.0007(0.0456)	−0.0103(0.0487)	95.0 %
$\sigma^2_{\nu_0}=1.96$	0.0077(0.3510)	−0.0942(0.4650)	0.0123(0.1994)	−0.0197(0.2684)	45.6 %
$\sigma^2_\delta=1$	0.9996(1.013)	−0.0795(0.2950)	1.0052(1.0095)	−0.0289(0.1760)	39.6 %

Table 4 Bias (RMSE) of the MLE and MME estimators based on 1000 Monte-Carlo simulations for the Berkson ME model with ME on both X and Y

True value	$n=100$		$n=300$		CP
	MLE	MME	MLE	MME	
$\beta_0=8$	0.0062(0.1595)	−0.0169(0.1595)	−0.0013(0.0939)	−0.0107(0.0977)	94.3 %
$\beta_1=2$	0.0023(0.0962)	−0.0295(0.0959)	−0.0012(0.0555)	−0.0080(0.0585)	95.2 %
$\sigma^2_{\nu_0}=1.96$	0.0155(0.3973)	0.0180(0.5260)	0.0050(0.2243)	0.0088(0.2881)	35.9 %
$\sigma^2_\delta=1$	1.9948(2.0098)	−0.1580(0.4084)	2.0024(2.0072)	−0.0485(0.2199)	45.0 %

Table 5 Bias (RMSE) of the MME estimators based on 1000 Monte-Carlo iterations for a misspecified ME model with ME on X, $n=100$

True value	Bias	RMSE
$\beta_0=8$	−0.0552	0.1740
$\beta_1=2$	−0.0238	0.0851
$\sigma^2_{\nu_0}=1.96$	0.3993	0.8383
$\sigma^2_\delta=1$	2.5306	2.6248

expected, the naive estimator for all the parameters (except $\sigma^2_{\nu_0}$) is more biased than MME. We also notice that the bias in MLE is persistent even when we increase the sample size to 300 while the bias in MME decreases. MLE has a smaller bias than MME in estimating $\sigma^2_{\nu_0}$. Theoretically, MLE of $\sigma^2_{\nu_0}$ is unbiased [32]. This is not very surprising, since MLE is using the strength of the full information on the distribution of X, δ and ν_0. Both estimators have smaller bias on $\sigma^2_{\nu_0}$ when the sample size increases. The large bias in the naive variance estimator of δ shows an overestimation of the variability of the model error term. This bias increases even more when MLE ignores the ME on both X and Y.

4.5.2 Case Two—Berkson ME

Tables 3 and 4 summarize the results for the Berkson case with either ME on X only (Table 3) or ME on both X and Y (Table 4). Although MME shows a larger finite sample bias in estimating β_0, β_1 and $\sigma^2_{\nu_0}$, as we theoretically demonstrated, MLE has a much larger bias in estimating σ^2_δ. The finite sample bias in MME reduces with increasing n, but this is not the case for MLE. This bias increases to a very large amount when we have ME on both X and Y. This indicates the large impact of ignoring ME on both covariate and response for the naive estimator. As for the classical model, the finite sample bias in MME is due to the minimum use of information regarding the distribution of the variables in the models. As previously stated, an additive classical ME is the most common model for ME on variables. As suggested from the results of the simulations, in this case, MME can provide much more reliable estimators for all the parameters of the model than the naive ML estimator especially when the sample size is large enough.

In Tables 1–4, we also added the coverage probabilities (CP) of the MM estimators for $n = 300$. The results show that the average coverage probability for the fixed effects is about 95 %. The coverage probability of the variance of the random effect varies between 35.9 % and 56.8 %.

4.5.3 Case Three—Misspecified ME

Table 5 shows that under the misspecified ME model for X, MME still provides quite satisfying estimators for fixed effects, even for a relatively small sample size. Although the estimators for the variance of the random effect $\sigma^2_{\nu_0}$ and the model error term σ^2_δ are biased, the results are encouraging because fixed effects are often of more interest. Considering that MME does not use any distributional assumptions on any of the random variables in the model, it still provides satisfactory estimators for most of the parameters of interest in real applications. The large biases in $\sigma^2_{\nu_0}$ and σ^2_δ can be explained by two factors. First, $\mathrm{var}(W) > \mathrm{var}(X)$ implied by the classical ME model is a wrong assumption (under the misspecified ME), when the true Berkson ME model requires that $\mathrm{var}(W) < \mathrm{var}(X)$. More generally, the ME model is a part of the full model, so if the ME model is misspecified, then the full model is misspecified. Second, V has a weaker correlation with X than W. So the estimates based on V are usually less accurate than those based on W.

Comparing Tables 3 and 5, one might find that MLE is a better estimator in the case of misspecification. MLE does not need the ME model and, for the Berkson model, it provides unbiased estimators for both random and fixed effects (see Sect. 3 and Buonaccorsi and Lin [11]), so it can be a better choice. If we know the distributions of all the variables in the model and the ME model is Berkson, but we do not know its specification, then the MLE could be a better option. Generally, however, ignoring ME leads to biased estimators even if we know the distributions of the variables.

5 Application—Raine: A Birth and Child Cohort Study

The Western Australian birth and child (Raine) cohort is an ongoing health research study in which pregnant women were recruited between 16 and 18 weeks gestation and their children followed from birth to 13 years. The LMEM was used to model the children's BMI growth trajectories in this study as a function of the gene FTO (fat mass- and obesity-associated) and, more particularly, the single-nucleotide polymorphism (SNP) rs9939609 in this gene following from the work of Abarin et al. [1]. The purpose of our study is to test for an interaction between this SNP and duration of breast feeding (BF) accounting for possible ME on BMI and BF. We studied a sample of 1096 children who were followed from birth to 13 years. The following model is considered for BMI at time $j = 0, 1, 2, 3, 8, 10, 13$ for an individual i:

$$\text{BMI} = \begin{cases} a_0^{\text{Inf}} + a_1^{\text{Inf}}(\text{Age} - 1.5) + \alpha_2^{\text{Inf}}(\text{Age} - 1.5)^2 + X\beta + \epsilon & \text{if Age} < 1.5 \\ a_0^{\text{Ch}} + a_1^{\text{Ch}}(\text{Age} - 1.5) + \alpha_2^{\text{Ch}}(\text{Age} - 1.5)^2 + X\beta + \epsilon & \text{if Age} \geq 1.5 \end{cases}$$

where X is a vector of fixed effects including the FTO SNP (coded with two dummy variables for the heterozygotes (TA) and homozygotes of the rare allele (AA)), BF, the interaction between the SNP and BF, the interaction between BF and age, and the interaction between the SNP and age. In the equation, Inf is used for Age < 1.5 as an infant, and Ch is used for Age ≥ 1.5 as a child. The coefficients a_1^{Inf}, a_1^{Ch} represent the mean acceleration/deceleration of BMI during infancy and childhood, and the index 0 stands for a random intercept and 1 for a random slope. In order to enforce continuity between the two time windows, we rewrite the model which enforces continuity at the breakpoint 1.5 as

$$\begin{aligned} Y = a_0 + a_1(\text{Age} - 1.5) + I[\text{Age} < 1.5]\alpha_2^{\text{Inf}}(\text{Age} - 1.5)^2 \\ + I[\text{Age} \geq 1.5]\alpha_2^{\text{Ch}}(\text{Age} - 1.5)^2 + X\beta + \epsilon, \end{aligned} \tag{23}$$

where I is an indicator function.

Since the BMI growth trajectories are not identical across individuals, we also included the random intercept ν_0 and random slop ν_1 in the model. These two random effects are assumed to follow a bivariate normal distribution with mean 0 and variance–covariance matrix

$$\Sigma_\nu = \begin{pmatrix} \sigma_{\nu_0}^2 & \sigma_{\nu_{01}} \\ \sigma_{\nu_{01}} & \sigma_{\nu_1}^2 \end{pmatrix}.$$

Since the average body mass index and age of puberty differs between boys and girls, we fitted separate models for males and females. We should mention that our data set was unbalanced, that is, we had unequal numbers of observations of each time point. More information regarding the Raine study and also the justification for this model selection can be found in Abarin et al. [1].

In most longitudinal research studies, when BMI at a certain age is collected, the variable of interest for BMI is actually the long term average value of BMI for the person in that year. The reason why the true and observed BMI differ is that weight has daily as well as seasonal variation. Moreover, since BMI only takes into consideration overall weight and height, it can cause an overestimation or underestimation of the true BMI.

Some epidemiological studies showed that self-reported information on duration of exclusive breast feeding tends to be biased [7, 29]. The main reason is that generally the duration of breast feeding is mixed with other kind of milks and solids which can mask the real impact of "exclusive" breast feeding (EXBF). In the modeling of BMI growth trajectories, our interest is therefore to propose a ME model for the duration of breast feeding (BF) considering EXBF as close to the true value. We select a classical model for the ME of BF as it seems there is more variability in the observed (BF) than the true value (EXBF) [29]. Another motivation is that BF measurements can be replicated. In some studies, such as measures of radiation exposure, replicates are not available. We considered BF as time invariant because it is observed once for every individual. The instrumental variable V that we use for the study is the minimum value of the age that women stopped breast feeding and the age at which mothers started to feed their babies with other kind of milks. Our study on the Raine data shows that V is related to BF according to

$$E(\mathrm{BF}|V) = 0.08 + 0.88V,$$

where U is independent from V and ε.

For the ME in the response, a classical model seems reasonable because, according to Carroll et al. [14], BMI is measured uniquely for an individual and it can also be replicated. All computations were done in R and the naive ML estimates are obtained from the `lmer` package. Since we had no prior information or validation data, we could not estimate σ_ξ^2, so σ_δ^2 is not identifiable either. The parameters σ_ε^2 and σ_u^2 were not of interest (nuisance), so we did not report any of these estimates in the tables of the results.

As we mentioned earlier, to estimate the working matrix for our unbalanced data, we divided our individuals into groups that had the same number of observations. Then we used the diagonal matrix form of the weighting matrix to compute H. Tables 6 and 7 show the estimates of parameters using both the naive MLE and the method of moment estimations as well as the standard errors of the estimates.

It appears that for most of the fixed effects related to age, the MME approach produces values close to the naive MLE approach. Wang et al. [32] showed that the naive estimator of the intercept for Gaussian data is asymptotically biased. Therefore, we can conclude that we have better estimators of the intercept in MME than the naive estimator. The rest of the effects seem to be quite different between these two approaches. Carroll et al. [14] showed that the naive estimator of the effect on the accurately measured covariate that is dependent on the error-prone covariate is biased. Thus, in all the interaction terms between BF and other covariates, we expect to have more accurate results in MME. It also appears that the naive MLE mostly yields smaller standard errors. With respect to the estimation of variabilities of the random effects, both approaches yield to similar results. Moreover, theoretically, we do not expect to have much difference between the naive estimator and MME.

Table 6 Estimates of fixed effects and the standard errors for MME and the naive MLE—Males

	Naive	SE	MME	SE
Intercept	16.61	0.14	16.92	0.36
$I\,(\text{Age} < 1.5)$TRUE	0.30	0.28	−0.1	0.42
$(\text{Age} - 1.5)\ (< 1.5)$	−2.80	0.95	−3.71	0.94
$(\text{Age} - 1.5)^2 (< 1.5)$	−3.51	0.51	−3.99	0.50
$(\text{Age} - 1.5)\ (\geq 1.5)$	−0.31	0.03	−0.36	0.05
$(\text{Age} - 1.5)^2 (\geq 1.5)$	0.05	0.001	0.06	0.004
BF	−0.02	0.01	−0.04	0.001
TA	−0.04	0.16	−0.21	0.19
AA	0.02	0.24	−0.37	0.64
(Age − 1.5):BF	−0.003	0.002	−0.005	0.009
BF:TA	0.002	0.01	0.02	0.02
BF:AA	0.005	0.02	0.04	0.10
(Age − 1.5):TA	0.12	0.03	0.12	0.01
(Age − 1.5):AA	0.10	0.04	0.11	0.01
σ_{ν_0}	0.89	–	0.71	0.29
σ_{ν_1}	0.30	–	0.31	0.002
$\sigma_{\nu_{01}}$	0.27	–	0.25	0.1

Table 7 Estimates of fixed effects and the standard errors for MME and the naive MLE—Females

	Naive	SE	MME	SE
Intercept	15.75	0.14	16.29	0.59
$I\,(\text{Age} < 1.5)$TRUE	0.67	0.29	0.41	0.26
$(\text{Age} - 1.5)\ (< 1.5)$	−1.11	0.99	−1.42	0.30
$(\text{Age} - 1.5)^2 (< 1.5)$	−2.34	0.54	−2.54	0.18
$(\text{Age} - 1.5)\ (\geq 1.5)$	−0.21	0.03	−0.31	0.09
$(\text{Age} - 1.5)^2 (\geq 1.5)$	0.05	0.001	0.06	0.005
BF	0.01	0.01	−0.02	0.09
TA	0.41	0.17	0.12	0.42
AA	0.005	0.24	0.10	0.45
(Age − 1.5):BF	−0.001	0.002	0.002	0.003
BF:TA	−0.02	0.01	0.01	0.11
BF:AA	0.002	0.02	−0.02	0.16
(Age − 1.5):TA	0.01	0.03	0.01	0.01
(Age − 1.5):AA	−0.01	0.04	−0.03	0.01
σ_{ν_0}	0.93	–	0.98	0.32
σ_{ν_1}	0.30	–	0.29	0.05
$\sigma_{\nu_{01}}$	0.21	–	0.2	0.09

6 Summary and Discussion

We defined a linear mixed effects model with measurement error on the fixed covariates and on the response for longitudinal data. We also proposed a semi-parametric methodology that does not rely on any assumption concerning the functional forms of the distributions of covariates, random effects, response and measurement errors. Using moment equations, we showed that all the parameters of the model can be estimated. We also applied our methodology to both classical and Berkson ME on covariates. Using simulation studies and an example motivated by a large birth and child cohort study from Australia (Raine), we investigated the finite sample performances of the estimators and showed the impact of measurement error on the covariates, and response on the estimation procedure. The results of both studies show that our procedure performs quite satisfactory especially for the fixed effects with measurement error (even when we have misspecification on the ME model). To the authors' knowledge, no other study has applied a semi-parametric method for longitudinal data with measurement error that does not need replicates or is computationally feasible. There are some methods that need replicates in order to estimate the model parameters, however, those for longitudinal data can be quite expensive. The method we proposed here is novel and needs only instrumental variables. It is, therefore, less restrictive than using replicates and it can be applied in many areas of study and research, including gene-environmental interaction studies. As we mentioned in the introduction, ME is usually ignored in these studies. We showed in our application to the Raine data that ignoring ME can cause some serious statistical problems. More specifically, we showed that the fixed environmental effects with ME (including the gene-environmental interaction) can be under/over estimated with large bias if we apply only the naive estimator. It is also noteworthy that if we study associations between genes and complex diseases, the naive estimator can cause false negatives in testing hypotheses.

Regarding the model assumptions in Sect. 2, we considered that δ_{ij} and also ξ_{ij} are mutually independent with constant variances. The first one leads to a "compound symmetric" error covariance matrix. Our methodology can be applied to a more general covariance matrix structure, however, since this structure was tested and found suitable for our analysis of the Raine data which we also used in this paper. Since BMI is measured independently for different individuals and different time points, it is reasonable to assume that the ME on BMI is independent between individuals as well as between observations, a common assumption in ME literature.

A typical Berkson or classical ME model has a random error term with mean 0. Such an assumption for the duration of breast feeding may not be realistic as the measurement error may not have mean 0. The ME in the covariate in the Raine data seems to come from inaccurate measurement of the duration of exclusive breast feeding. Moreover, women who fed their babies with other liquids or solids (at the same time as breast feeding), tend to overestimate the duration. In this case, since we do not expect the same amount of underestimation, the measurement error might not have a zero mean. In the past decade, there have been studies on flexible ME [20, 27], however, all these focused on the modeling of different variables of interest (mostly energy or calory intake) in food frequency questionnaires. To the authors' knowledge, there are no studies on the modeling of duration of breast feeding. Further studies are also needed to find a better model for ME on duration of breast feeding. In the future, we are interested in studying more flexible forms of ME models, as well as methods for ME on categorical variables. The latter topic is important in genetic association studies between complex diseases and genotypes when there is a significant genotyping error.

Acknowledgements We thank the referees for their comments and helpful suggestions for manuscript changes. These have improved both the content and clarity of the manuscript. We would also like to thank Wei Ang, Nicole Warrington, Julie Marsh and Louise Mckenzie, for their help in preparing the data set and choosing the variables of the model in the application, and the RAINE study (http://www.rainestudy.org.au/) for providing the data for analysis. Financial support from "The Alva Foundation" and "Samuel Lunenfeld Research Institute—New Opportunities Funds" and MITACS are gratefully acknowledged.

References

1. Abarin T, Wu Y, Warrington N, Pennell C, Briollais L (2011) The impact of breastfeeding on FTO-related BMI growth trajectories: an application to the RAINE birth and child cohort study. Int J Epidemiol (to appear)
2. Abarin T, Wang L (2010) Instrumental variable approach to covariate measurement error in generalized linear models. Ann Inst Stat Math. doi:10.1007/s10463-010-0319-0
3. Abarin T, Wang L (2009) Second-order least squares estimation of censored regression models. J Stat Plan Inference 139:125–135
4. Abarin T, Wang L (2006) Comparison of GMM with second-order least squares estimation in nonlinear models. Far East J Theor Stat 20(2):179–196
5. Berkson J (1950) Are there two regressions? J Am Stat Assoc 45:164–180
6. Biggs R, et al (2009) Spurious certainty: how ignoring measurement error and environmental heterogeneity may contribute to environmental controversies. Bioscience 59(1):65–76
7. Bland RM, et al (2003) Maternal recall of exclusive breast feeding duration. Arch Dis Child 88(9):778–783
8. Buonaccorsi J (2010) Measurement error: models, methods, and applications. Chapman & Hall, London
9. Buonaccorsi J (1996) Measurement error in the response in the general linear model. J Am Stat Assoc 91:633–642
10. Buonaccorsi J, Tosteson T (1993) Correcting for nonlinear measurement errors in the dependent variable in the general linear model. Commun Stat, Theory Methods 22:2687–2702
11. Buonaccorsi J, Lin Ch (2002) Berkson measurement error in designed repeated measures studies with random coefficients. J Stat Plan Inference 104:53–72
12. Buonaccorsi J, et al (2000) Estimation in longitudinal random effects models with measurement error. Stat Sin 10:885–903
13. Burton PR, et al (2009) Size matters: just how big is BIG? Quantifying realistic sample size requirements for human genome epidemiology. Int J Epidemiol 38:263–273
14. Carroll RJ, et al (2006) Measurement error in nonlinear models: a modern perspective, 2nd edn. Chapman & Hall, London
15. Carroll RJ, Wand MP (1991) Semiparametric estimation in logistic measurement error models. J R Stat Soc B 53:573–585
16. Faith MS, et al (2004) Parental feeding attitudes and styles and child body mass index: prospective analysis of a gene-environment interaction. Pediatrics 114(4):e429–e436
17. Fuller AW (1987) Measurement error models. Wiley, New York
18. Joseph ML, Carriquiry A (2010) A measurement error approach to assess the association between dietary diversity, nutrient intake, and mean probability of adequacy. J Nutr 140(11):2094S–2101S
19. Kamarainen AM, et al (2008) Zooplankton and the total phosphorus—chlorophyll a relationship: hierarchical Bayesian analysis of measurement error. Can J Fish Aquat Sci 65(12):2644–2655
20. Kipnis V, et al (2003) Structure of dietary measurement error: results of the OPEN biomarker study. Am J Epidemiol 158(1):14–21

21. Laird NM, Ware JH (1982) Random-effects models for longitudinal data. Biometrics 38:963–974
22. Li H (2005) A simulation study of the second-order least squares estimators for nonlinear mixed effects models. Master's thesis, University of Manitoba
23. Li H, Wang L (2011) Second-order least squares estimation in linear mixed models. University of Manitoba. Comm Statist Theory Methods doi:10.1080/s03610926.2011.601837
24. Marchand L, Wilkens L (2008) Design considerations for genomic association studies: importance of gene-environment interactions. Cancer Epidemiol Biomark Prev 17:263–267
25. Pan W, et al (2009) Semiparametric transition measurement error models for longitudinal data. Biometrics 65:728–736
26. Parsons TJ, et al (2001) Fetal and early life growth and body mass index from birth to early adulthood in 1958 British cohort: longitudinal study. Br Med J 323:1331–1335
27. Prentice RL, et al (2002) Research strategies and the use of nutriet biomarkers in studies of diet and chronic disease. Public Health Nutr 5:977–984
28. Rampersaud E, et al (2008) Physical activity and the association of common FTO gene variants with body mass index and obesity. Arch Intern Med 168(16):1791–1797
29. Rios E, et al (1992) Accuracy of mothers' responses to questions about breast-feeding practices. Food Nutr Bull 14(2):115–118
30. Schennach MS (2007) Instrumental variable estimation of nonlinear errors-in-variables models. Econometrica 75:201–239
31. Thomas D (2010) Methods for investigating gene-environment interactions in candidate pathway and genome-wide association studies. Annu Rev Public Health 31:21–36
32. Tosteson TD, et al (1998) Covariate measurement error and the estimation of random effect parameters in a mixed model for longitudinal data. Stat Med 8:1139–1147
33. Tsiatis T, Davidian M (2001) A semiparametric estimator for the proportional hazards model with longitudinal covariates measured with error. Biometrika 88(2):447–458
34. Vimaleswaran KS, et al (2009) Physical activity attenuates the body mass index–increasing influence of genetic variation in the FTO gene. Am J Clin Nutr 90(2):425–428
35. Wang L (2004) Estimation of nonlinear models with Berkson measurement errors. Ann Stat 32:2559–2579
36. Wang L, Hsiao C (2010) Method of moments estimation and identifiability of nonlinear semiparametric errors-in-variables models. J Econom 165:30–44. doi:10.1016/j.jeconom.2011.05.004
37. Wang N, et al (1998) Bias analysis and SIMEX approach in generalized mixed measurement error models. J Am Stat Assoc 93:249–261
38. Wang N, et al (1999) A bias correction regression calibration approach in generalized linear mixed measurement error model. Commun Stat, Theory Methods 28:217–232
39. Wong MY, et al (2003) The detection of gene-environment interaction for continuous traits: should we deal with measurement error by bigger studies or better measurement? Int J Epidemiol 32:51–57