J. Statist. Res. Iran **16** (2019): 255–286
DOI: ?

Likelihood Inference in the Random Effects Logistic Regression Model with Response Misclassification and Covariate Subject to Measurement Error

Maryam Ahangari†, Mousa Golalizadeh*,† and
Zahra Rezaei Ghahroodi‡

† Tarbiat Modares University
‡ University of Tehran

Received: 11/26/2020 Approved: 1/31/2021

Abstract. Generalized linear mixed models (GLMMs) are common methods for the analysis of clustered data. In many longitudinal and hierarchical epidemiological frameworks, accurate measurements of variables are invalid or expensive to be obtained and there might be situations that both the response and covariate variables are likely to be mismeasured. Insensitivity of errors in either covariate or response variable is, not always plausible. With nonlinear regression models for the outcome process, classification errors for binary responses and measurement error in covariates basically needs to be accounted for in order to make conclusive inferences. In this article, we provide an approach to simultaneously adjust for non-differential misclassification in the correlated binary response and classical measurement error in the covariates, using the multivariate Gauss-Hermite quadrature technique for the approximation of the likelihood function. Simulation studies are then conducted to inform the effects of correcting for measurement error and misclassification on the estimation of regression parameters. The application of the multivariate Gauss-Hermite quadrature method in the conjunction of measurement error and misclassification problems is further highlighted with

* Corresponding author

real-world data based on a multilevel study of contraceptive methods used by women in Bangladesh.

Keywords. Measurement Error; Binary Response; Multivariate Gauss-Hermite Quadrature; Random Effects Logistic Regression Model; Misclassification.
MSC 2010: 62J12.

1 Introduction

The prime intention of longitudinal and multilevel models is to form a pattern for the relationship between the response variable and a set of covariates by incorporating the correlation between repeated measurements or clustered structure. Conceding that the response variable is non-normally distributed, generalized linear mixed models (GLMMs) are extensively adopted, considering a nonlinear link between the mean of the response variable with a set of predictors. In this manner, by adding random effects, the correlation between responses can be modeled appropriately (McCullach et al., 2008). In the estimation process, the application of maximum likelihood (ML) method is limited in GLMMs, which is particularly as a result of non-normality of the conditional distribution of the response variable given the random effects, hereof the likelihood function may not have a closed form. To avoid computational problems surrounding the ML method, Monte Carlo EM (MCEM) approach was proposed for the analysis of binary responses (McCullach, 1994). Moreover, other numerical algorithms including Laplace approximation or Gaussian quadrature have also been applied to suppress high dimensional integrals in the likelihood function. For more details, researchers can consult Wu (2009) and Stroup (2012). Also, Tanner (1993) and Diggle et al. (1994) are recommended on Bayesian aspect to approximate the integrals.

In many applications, some of the predictors of GLMMs may not be measured preciesly, so measurement error in continuous covariates is a challenging problem in the analysis of multilevel-structured data with discrete responses. General strategies of handling covariate mismeasurement effects in estimating the parameters corresponding to correlated data are regression calibration (Wang et al. 1999), simulation extrapolation (Wang et al. 1998) and the bootstrap approach (Buonaccorsi, 2018). The Bayesian method has also been applied in generalized linear mixed measurement error models due to its computational conveniency (Carroll et al., 2006). Data cloning approach proposed by Lele et al. (2007), has also been applied to generalized linear mixed measurement error models (Torabi, 2013). An alternative to the aforementioned methods is using the Monte Carlo scheme. Recently, Xie et al. (2017) have shown that the Monte Carlo Newton-Raphson (MCNR) method gives accurate estimates in the logistic regression model setting for analyzing longitudinal biomarker data, accounting for left-censoring and covariate measurement error.

As a special source of error, the misclassification of categorical response variables in the regression analysis is an important topic in considering the validity of paramenter estimation and also making further statistical inferences. For this reason, correcting for misclassification has been a matter for researchers in order to get trustworthy point estimates as well as valid confidence intervals. For more details, researches by Buonnacorsi (2010) is suggested. In relation with response misclassification, Magder and Hughes (1997), designed a logistic regression model and found the ML estimates using an expectation-maximization algorithm. Moreover in a generalized linear model context, Neuhaus (1999) determined the fact that overlooking misclassified responses results in highly biased estimates, along with reduced statistical efficiency. In situations with differential classification error in the response variable, Gerlach and Stamey (2007), have applied a Bayesian framework for the estimation process. In such a case, Carroll et al. (2006) have utilized likelihood-based methods. Besides, Lyles et al. (2011) have implemented the ML approach along with internal validation data in order to adjust for misclassified binary outcomes.
In the clustered and longitudinal binary data framework, Neuhaus (2002) quantified the magnitude of asymptotic bias in estimating the logistic regression parameters with response subject to classification error. Paulino et al. (2003) have also proposed a Bayesian framework for the analysis of binary responses subject to response classification error using a random effect logistic regression model. Additionally, Tang et al. (2015) focused on longitudinal validation data-based studies with repeatedly measured error-prone responses. They assumed differential misclassification mechanism and utilized the ML approach to adjust for misclassification in the binary response. As previously mentioned, failure to incorporate covariate measurement error or response misclassification will inevitably cause invalid biased estimates and threats the statistical inferences. As noted by Cheng et al. (2010), measurement error in covariates together with misclassification in the outcomes

occur frequently in epidemiological studies, and failure to account for these two major sources of error, results in potentially biased and inaccurate estimates. They reported the fact that by ignoring measurement error with misclassified logistic regression data, the statistical power will be miscalculated. For this purpose, they adopted a Bayesian framework to estimate the regression parameters. Against this background, there are insufficiency of researches covering inference procedures for handling response misclassification together with covariate measurement error. Yi (2016) has explored methods for handling error in response and covariates simultaneously, regarding to both univariate and correlated data. Roy (2012) has further surveyed these two sources of errors in the longitudinal data structure with the intention of correcting for response classification errors and covariate measurement error, by utilizing the Monte Carlo Markov Chain (MCMC) computational technique.

In this paper, we develop the multivariate Gauss-Hermite quadrature method for approximating the intractable integrals of the likelihood function for multilevel data involving binary correlated response components in conjunction with non-differential misclassified responses and covariate measurement error. According to simulation outcomes, this technique of approximating the likelihood function can grant accurate estimates of the fixed effects of regression parameters and variance components of random effects and the measurement error distribution, as well as the sensitivity and specificity values. At an early stage, in Section 2, we will outline generalized linear mixed misclassification and measurement error models for binary outcomes with response subject to classification errors and covariates expose to measurement error. In Section 3, we describe the procedure by which the likelihood function can be assessed in the presence of mixed models in tandem with misclassified responses and covariates subject to classical measurement error. Afterwards, we represent the way by which the multivariate Gauss-Hermite quadrature procedure can be implemented to approximate the likelihood function adjusting for response misclassification and covariate measurement error. In Section 4, the performance of the multivariate Gauss-Hermite quadrature method to correct for induced bias will be assessed in a simulation based sensitivity analysis, accounting for various reliability ratios and different scopes of sensitivity and specificity measures, concurrently. We will examine the efficiency and competence of our proposed approach to correct for misclassified binary responses together with measurement error in the covariates and evaluate how misspecifying the measurement error distribution or not adjusting for classification error in the response variable or overlooking both of these errors might result in high biases and misinterpretations of statistical inferences. In Section 5, the improvement of the proposed method will be evaluated by analyzing a real dataset depending on a multilevel study on contraceptive methods utilized in Bangladesh. Concluding remarks are provided in the last section.

2 The Generalized Linear Mixed Misclassification and Measurement Error Model, Basics and Notation

Suppose $Y_i = (Y_{i1}, \cdots, Y_{in_i})^T$ denotes the observed outcomes for the ith subject, $1 \leq i \leq m$, where m is the total number of independent individuals, and n_i is the number of observations for individual i. Let Y_i be a misclassification-prone response variable that follows a generalized linear mixed model with random intercept for each individual. The vector of model covariates $X_i = (X_{i1}^T, \cdots, X_{in_i}^T)^T$ and $Z_i = (Z_{i1}^T, \cdots, Z_{in_i}^T)^T$ are assumed to be time-varying. Let Y_{ij} be the response variable for subject i at the jth occasion measurement ($j = 1, \cdots, n_i$), X_{ij} as the vector of true covariates which are error-prone and unobserved and hence latent, and Z_{ij} be the vector of error-free covariate. Furthermore, let W_{ij} and Y_{ij}^* be the palpable and surrogate measurements of X_{ij} and Y_{ij}, respectively, that are observed in the absence of X_{ij} and Y_{ij}.

2.1 The Outcome Process, Measurement Error Models and Classification Error Probabilities

Let Y_{ij} be a binary response variable for subject i at occasion j. Assume that the binary outcome is related with the set of covariates through the random effects logistic regression model as

$$logit[P(Y_{ij} = 1|x_{ij}, z_{ij}, \tau_i)] = \beta_0 + \beta_x x_{ij} + \beta_z z_{ij} + \tau_i, \quad i = 1, \cdots, m, \quad j = 1, \cdots, n_i. \quad (1)$$

In (1), $\beta = (\beta_0, \beta_x, \beta_z)^T$ is the vector of fixed parameters. We assume that the random effects τ_i, ($i = 1, \cdots, m$) are independent from each other, and also independent from the error-prone covariates X_{ij}. In this article, it is assumed that $\tau_i \sim N(0, \sigma_\tau^2)$, where σ_τ^2 is unknown and assumed to be constant for all

subjects.
In (1), the covariate X_{ij} is not directly observable, and W_{ij} is the only available measurement in the absence of X_{ij}, so W_{ij} can be considered as the surrogate variable for X_{ij}. In the following, it is assumed that W_{ij} is a composite of X_{ij} plus some error, expressly classical measurement error, in which the true but latent variable is measured with additive error, usually assumed to have constant variance, i.e.

$$\boldsymbol{W_{ij}|x_{ij} = x_{ij} + e_{ij}.} \tag{2}$$

where $\boldsymbol{e_{ij}}$ is the error term such that $E(e_{ij}|x_{ij}) = 0$ and $Var(e_{ij}|x_{ij}) = \sigma_e^2$, when W_{ij} and X_{ij} are scalar. It is concluded that $E(W_{ij}|x_{ij}) = x_{ij}$, so W_{ij} is unbiased for the unobserved x_{ij}. We also consider that the measurement error in W_{ij} is non-differential. This means that Y_{ij} is conditionally independent of W_{ij}, given x_{ij} and if X_{ij} is observed, the surrogate variable W_{ij} adds nothing to the predictor of Y_{ij}, i.e.;

$$P(Y_{ij} = 1|x_{ij}, w_{ij}, z_{ij}, \tau_i) = P(Y_{ij} = 1|x_{ij}, z_{ij}, \tau_i). \tag{3}$$

In the following, it is assumed that we have homoscedastic measurement error, which refers to the case where the variance of W_{ij} given x_{ij} is constant. With random covariate X_i subject to error, we assume distributional assumption on error-prone covariate X_i, i.e. structural approach in which covariate X_i has a normal distribution with mean μ_x and variance Σ_x, and is independent of τ_i. We choose a fully structural case for the falliable covariate with X_i assumed i.i.d with

$$X_i = (X_{i1}^T, \cdots, X_{in_i}^T)^T \sim N(\mu_x, \Sigma_x) \quad i = 1, \cdots, m,$$

where

$$\Sigma_x = \begin{pmatrix} \sigma_{x1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{x2}^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \cdots & \sigma_{xn_i}^2 \end{pmatrix}.$$

The measurement error in the multivariate case, also has a normal distribution with the following structure:

$$e_i \sim N(0_{n_i}, \sigma_e^2 I_{n_i}).$$

Suppose that Y_{ij} is also subject to misclassification, and Y_{ij}^* is an observed value of Y_{ij}. Particularly, let

$$P(Y_{ij}^* = 1|Y_{ij} = 0, x_{ij}, z_{ij}, \tau_i) = P(Y_{ij}^* = 1|Y_{ij} = 0) = \gamma_{01} \tag{4}$$

and

$$P(Y_{ij}^* = 0|Y_{ij} = 1, x_{ij}, z_{ij}, \tau_i) = P(Y_{ij}^* = 0|Y_{ij} = 1) = \gamma_{10}, \tag{5}$$

denote the simple case of misclassification probabilities. The probability $1 - \gamma_{10}$ is often called the sensitivity of the measurement Y_{ij}^*, and $1 - \gamma_{01}$ is called the specificity. In (4) and (5), γ_{01} and γ_{10} are considered to be unrecognized, nonnegative classification error probabilities that are constant and independent of other covariates in the model. This notification is widely referred to as 'non-differential misclassification'. According to the notations given so far, the conditional probability for the observed response given the true values of covariates and the random component is provided as follows:

$$\begin{aligned} P_{ij}^* &= P(Y_{ij}^* = 1|x_{ij}, z_{ij}, \tau_i) \\ &= P(Y_{ij}^* = 1|Y_{ij} = 1, x_{ij}, z_{ij}, \tau_i)P(Y_{ij} = 1|x_{ij}, z_{ij}, \tau_i) \\ &+P(Y_{ij}^* = 1|Y_{ij} = 0, x_{ij}, z_{ij}, \tau_i)P(Y_{ij} = 0|x_{ij}, z_{ij}, \tau_i) \\ &= (1 - \gamma_{10})P_{ij} + \gamma_{01}(1 - P_{ij}) = \gamma_{01} + (1 - \gamma_{10} - \gamma_{01})P_{ij}, \end{aligned} \tag{6}$$

where the quantity P_{ij} is described as

$$P_{ij} = P(Y_{ij} = 1|x_{ij}, z_{ij}, \tau_i) = \frac{\exp(\beta_0 + \beta_x x_{ij} + \beta_z z_{ij} + \tau_i)}{1 + \exp(\beta_0 + \beta_x x_{ij} + \beta_z z_{ij} + \tau_i)}. \tag{7}$$

Furthermore, let us consider W_{ij} as a surrogate for X_{ij} with classical additive measurement error as (2), where $e_{ij} \sim N(0, \sigma_e^2)$. Then, the conditional distribution of $W_{ij}|x_{ij}$ can be written as $W_{ij}|x_{ij} \sim N(x_{ij}, \sigma_e^2)$. Moreover, τ_i follows a normal density with mean zero and variance σ_τ^2. So the probability density function ($p.d.f$) for random effects in the single setting can be written as follows:

$$f(\tau_i; \sigma_\tau^2) = (2\pi\sigma_\tau^2)^{-\frac{1}{2}} \exp(-\frac{1}{2}\frac{\tau_i^2}{\sigma_\tau^2}). \tag{8}$$

It is important to mention that under non-differential measurement error mechanism, the conditional probability for the observed data may be written as:

$$\begin{aligned}
P(Y_{ij}^* = 1|w_{ij}, z_{ij}, \tau_i) &= \int P(Y_{ij}^* = 1|x_{ij}, w_{ij}, z_{ij}, \tau_i; \beta) f(w_{ij}|x_{ij}; \sigma_e^2) f(x_i; \mu_x, \Sigma_x) dx_{ij} \\
&= \int P(Y_{ij}^* = 1|x_{ij}, z_{ij}, \tau_i; \beta) f(w_{ij}|x_{ij}; \sigma_e^2) f(x_i; \mu_x, \Sigma_x) dx_{ij} \\
&= \int [\gamma_{01} + (1 - \gamma_{10} - \gamma_{01}) \\
&\quad \times P(Y_{ij} = 1|x_{ij}, z_{ij}, \tau_i; \beta)] f(w_{ij}|x_{ij}; \sigma_e^2) f(x_i; \mu_x, \Sigma_x) dx_{ij}
\end{aligned}$$

For more details, researchers can consult Buonaccorsi (2010) and Yi (2016). In the ensuing article, we manage a sensitivity analysis by supplying a scope of possible values for the classification error and measurement error parameters under the assumption of non-differential misclassification and measurement error mechanism. Next, we proceed to improve the model via correcting the measurement error along with classification error. It should be pointed out that to avoid complexity issues come to pass in estimating the parameters of interest in a simulation step, we will treat the mean together with the covariance matrix of the error-prone variable X_i to be fixed and develop a trajectory to correct for the induced bias in parameter estimation stage.

3 Likelihood Inference under both Misclassification and Measurement Error

This Section provides the way of constructing the likelihood function based on distributional assumptions recalled earlier. The likelihood formulation for the generalized linear mixed misclassification and measurement error model under the non-differential error structure involves up to four different parts: the model for Y_{ij}^* given x_{ij} for each of j (specific) occasions, the model for measurement error, the model for X_{ij} in the structural setting, and the model for the random effects. Next, we show how to build the likelihood function and then describe its proper approximation.

3.1 Likelihood Function arrangements and measurement error pattern

In this Section, our main focus lies in the case in which a parametric likelihood procedure is considered that allows for classification error in the binary responses and measurement error in the covariates in a random effects logistic regression model. According to (1) and (2), if $\theta = (\beta, \sigma_\tau^2, \sigma_e^2)^T$ is a vector of associated parameters along with the vector of misclassification rates $\gamma = (\gamma_{01}, \gamma_{10})^T$, for the ith subject at occasion j, given the observed data $(y_{ij}^*, w_{ij}, z_{ij})$ for $i = 1, \cdots, m$ and $j = 1, \cdots, n_i$, the likelihood function that simultaneously incorporates the measurement error process in covariates and the classification errors in the binary response, can be outlined as follows:

$$\begin{aligned} L(\theta; y^*, w) &= \prod_{i=1}^{m} \int_R \prod_{j=1}^{n_i} f(y_{ij}^*|w_{ij}, z_{ij}, \tau_i; \beta) f(\tau_i; \sigma_\tau^2) d\tau_i \\ &= \prod_{i=1}^{m} \int_R \int_{R^{n_i}} \prod_{j=1}^{n_i} f(y_{ij}^*|x_{ij}, w_{ij}, z_{ij}, \tau_i; \beta) f(w_{ij}|x_{ij}; \sigma_e^2) \\ &\quad \times f(x_i; \mu_x, \sigma_{xj}^2) f(\tau_i; \sigma_\tau^2) dx_{ij} d\tau_i \\ &= \prod_{i=1}^{m} \int_R \int_{R^{n_i}} \prod_{j=1}^{n_i} f(y_{ij}^*|x_{ij}, z_{ij}, \tau_i; \beta) f(w_{ij}|x_{ij}; \sigma_e^2) \\ &\quad \times f(x_i; \mu_x, \sigma_{xj}^2) f(\tau_i; \sigma_\tau^2) dx_{ij} d\tau_i, \end{aligned} \tag{9}$$

where $f(y_{ij}^*|x_{ij}, w_{ij}, z_{ij}, \tau_i; \beta)$ equaling to $f(y_{ij}^*|x_{ij}, z_{ij}, \tau_i; \beta)$ due to the assumption of non-differential measurement error, has a Bernoulli distribution given the random effects. For more details, referring to the Yi (2016) is recommended. This density function is defined as follows:

$$f(y_i^*|x_i, z_i, \tau_i; \beta) = \prod_{j=1}^{n_i} (P_{ij}^*)^{y_{ij}^*} (1 - P_{ij}^*)^{1 - y_{ij}^*}, \tag{10}$$

wher the quantity P_{ij}^* is a pointwise probability which is determined in (6). The likelihood function conveyed by (9), reflects the fact that the occurence of both classification error in the response and measurement error in the covariate typically has more intricate consequences in modifying the arrangement of the likelihood function when compared with situations with merely

misclassification in the response or measurement error in the covariate. As a result, the likelihood function (9), which encompasses both misclassification in the response variable and measurement error in the covariate, can be considered as the likelihood function corresponding to our proposed methodology. The suggested approach will be titled as the hybrid method during the succeeding sections. With the intention of having a comparative study about the results obtained from the hybrid method, we will contemplate three other scenarios, subsequently.
In the first scenario, we suppose that neither misclassification in the response variable nor the classical additive measurement error in the covariate has been covered in the estimation process. This model which naively disregards the difference between observed and true values of the binary response and the covariates is nominated as the 'naive' model. The likelihood function for this method is as follows:

$$L(\theta; y, w) = \prod_{i=1}^{m} \int_{R} \prod_{j=1}^{n_i} f(y_{ij}|w_{ij}, z_{ij}, \tau_i; \beta) f(\tau_i; \sigma_\tau^2) d\tau_i. \tag{11}$$

Another strategy is the one that corrects for misclassification in the response, uniquely. In this approach, the measurement error in the covariate is misspecified in the estimation procedure and it is assumed that covariates are precisely measured, which is denoted as the 'misclassification adjusted' method. The intention of offering this methodology is to scrutinize the effects of ignoring the measurement error in estimating the parameters of interest. In such a case, assessing bias or standard error results, can provide a comprehension for the situation in which the covariate measurement errors are not of consequence, while setting a substructure for misclassification correction method. The likelihood function corresponding to this methodology can be determined as:

$$L(\theta; y^*, w) = \prod_{i=1}^{m} \int_{R} \prod_{j=1}^{n_i} f(y^*_{ij}|w_{ij}, z_{ij}, \tau_i; \beta) f(\tau_i; \sigma_\tau^2) d\tau_i. \tag{12}$$

Alternative method that is also judged during this article is the measurement error modified approach. In this setting, it is assumed that the true values of the covariates are not observed directly and hence they are measured with classical additive measurement error. For this specific case, only the information on the proxy values for covariates are employed, though assuming the response variable is properly measured, i.e. treating the rsponse as error-free. This approach is hereafter denoted ‘measurement error adjusted’ method. Taking into account this approach, we aim to investigate the debilitations and fadings caused in the estimation of regression parameters due to the assumption of embracing the covariate measurement error but dismissing the response misclassification. The likelihood function in this approach is defined as:

$$L(\theta; y, w) = \prod_{i=1}^{m} \int_{R} \int_{R^{n_i}} \prod_{j=1}^{n_i} f(y_{ij}|x_{ij}, z_{ij}, \tau_i; \beta) f(w_{ij}|x_{ij}; \sigma_e^2) f(x_i; \mu_x, \sigma_{xj}^2) \times f(\tau_i; \sigma_\tau^2) dx_{ij} d\tau_i. \quad (13)$$

Unfavorably, the integrals in equations (9), (11), (12) and (13) take place in dimensions that their values depend on the random effects along with or without the latent true covariate X_{ij} on n_i occasions. Consequently, the observed data likelihood function for this model cannot be expressed in a tractable manner. This means that finding the ML estimates for the parameters will not be straightforward in application. For this case, standard numerical integration methods such as Multivariate Gauss-Hermite quadrature (MGHQ) methods can be imposed to approximate the integral (Jaeckel, 2005).

In a simple one dimensional setting, this means that the likelihood function is approximated using the Gauss-Hermite quadrature (GHQ) method and optimized with common approaches. The GHQ technique approximates the integral by a weighted sum of the integrand which is evaluated at a number of quadrature points in the domain of the integration function (Agresti, 2002). The location of the quadrature points and weights depend on the integrand and also on the domain of the integration function. For the integrals where the domain of the integration is $(-\infty, +\infty)$ or the entire real line, and the integrand is the product of a specific function with a normal density, the locations of the quadrature points are the solutions to the Hermite polynomial function (Skrondal and Rabe-Hesketh, 2004, Molenberghs and Verbeke, 2006).

As an illustration, for a univariate integral

$$I(f) = \int_{-\infty}^{\infty} f(x) \exp(-x^2) dx, \quad (14)$$

where $f(x)$ is an integrable function on the range $(-\infty, +\infty)$, the GHQ approximation technique, expresses the integral (14) in the format of a summation

$$I(f; \{x_i\}_{i=1}^q, \{g_i\}_{i=1}^q) = \sum_{i=1}^{q} g_i f(x_i), \tag{15}$$

where $\{x_i\}_{i=1}^q$ are called the nodes or quadrature points, which are the roots of the Hermite polynomial having degree q, i.e. $H_q(x)$, and the weights $\{g_i\}_{i=1}^q$ depend on q (the number of quadrature points) and the Hermite polynomial with degree $q-1$, say $H_{q-1}(x)$, evaluated at the ith node x_i by

$$g_i = \frac{2^{q-1} q! \sqrt{\pi}}{q^2 [H_{q-1}(x_i)]^2} \quad i = 1, \cdots, q.$$

The nodes $\{x_i\}_{i=1}^q$ are symmetric about zero (Pan and Thompson, 2003). The approximated likelihood function will then be maximized through invokong well-known standard algorithms such as the method proposed by Nelder and Mead (1965) or Newton-Raphson, yielding ML estimates for the parameters. To obtain the Monte Carlo estimates of the variance components for the ML estimates, one can calculate approximated observed information matrix.

3.2 Approximation of the Log-Likelihood Function

According to the likelihood function for the generalized linear mixed misclassification and measurement error model (9), the log-likelihood function based on the observed data incorporating the measurement error distribution and the classification errors can be written as follows:

$$l(\theta; y^*, w, z) = \sum_{i=1}^{m} \log[\int_R \int_{R^{n_i}} \prod_{j=1}^{n_i} f(y_{ij}^* | x_{ij}, z_{ij}, \tau_i; \beta) f(w_{ij} | x_{ij}; \sigma_e^2) \\ \times f(x_i; \mu_x, \sigma_{xj}^2) f(\tau_i; \sigma_\tau^2) dx_{ij} d\tau_i]. \tag{16}$$

Furthermore, the likelihood function for the observed data, misspecifying the misclassified responses and classical additive measurement error in the covariate X_i is

$$l(\theta; y, w, z) = \sum_{i=1}^{m} \log[\int_R \prod_{j=1}^{n_i} f(y_{ij}|w_{ij}, z_{ij}, \tau_i; \beta) f(\tau_i; \sigma_\tau^2) d\tau_i]. \quad (17)$$

For the approximation of the integral (16), multivariate forms of the Gauss-Hermite quadrature approximation method is needed to handle both unobserved covariate x_{ij} at n_i occasions and the random effects τ_i. Due to the fact that both x_{ij} and τ_i are independent from each other, the product of their density functions can be defined as a multivariate normal distribution. Multivariate Gauss-Hermite quadrature method (MGHQ) uses the idea of univariate GHQ rule for each coordinate of the integrated variables. For the approximation of the integrals in (16) which are $(n_i + 1)$ dimensional, a matrix of GHQ nodes are produced for latent variables X_{ij} at occasion j and also for the random effects τ_i according to the mean vector and variance-covariance matrix of the multivariate normal distribution $f(x_i, \tau_i; \sigma_{xj}^2, \sigma_\tau^2)$, which the integral is calculated. Furthermore, a vector of weights is calculated in the replacement with the multivariate normal distribution defined.

For the integrals in (16), the approximation process will be done by choosing sets of GHQ nodes

$$\{x_{kj} = (x_{k_1 j}^{(1)}, x_{k_2 j}^{(2)}, \cdots, x_{k_{(n_i+1)} j}^{(n_i+1)})' : \\ 1 \le k_1 \le q_1; 1 \le k_2 \le q_2; \cdots; 1 \le k_{(n_i+1)} \le q_{(n_i+1)}\}$$

and

$$\{\tau_k = (\tau_{k_1}^{(1)}, \tau_{k_2}^{(2)}, \cdots, \tau_{k_{(n_i+1)}}^{(n_i+1)})' : \\ 1 \le k_1 \le q_1; 1 \le k_2 \le q_2; \cdots; 1 \le k_{(n_i+1)} \le q_{(n_i+1)}\}$$

and weights

$$\{g_k = (g_{k_1}^{(1)}, g_{k_2}^{(2)}, \cdots, g_{k_{(n_i+1)}}^{(n_i+1)})' : \\ 1 \le k_1 \le q_1; 1 \le k_2 \le q_2; \cdots; 1 \le k_{(n_i+1)} \le q_{(n_i+1)}\}.$$

In this setting, k refers to the indices $k_1, k_2, \cdots, k_{(n_i+1)}$, each index shows different sets of quadrature points assigned, in order to approximate each

of (n_i+1) integrals (Pan and Thompson, 2003). Moreover, $x_{k_tj}^{(t)}$ and $\tau_{k_t}^{(t)}$ $(1 \leq t \leq (n_i+1))$ are the t-th root of the multivariate hermite polynomial having degree t, i.e. $H_{q_t}(x_j, \tau)$ at occasion j $(j = 1, \cdots, n_i)$.
It is important to notify that the total number of the integration quadrature points is $q = q_1 q_2 \cdots q_{(n_i+1)}$. But, if $q_1 = q_2 = \cdots = q_{(n_i+1)} = q_0$, i.e. the number of quadrature points for each dimension are assumed equal, then $q = q_0^{(n_i+1)}$. Adequate approximation usually needs larger number of grids for standard errors. It is recomended to increase the number of grids basically until the changes are negligible and inconsequential in both estimates and standard errors.
Thereupon, the MGHQ approximation to the log-likelihood function (16) will have the following form:

$$l(\theta; y^*, w, z) = \sum_{i=1}^{m} \log[\sum_{k_1=1}^{q_1} \sum_{k_2=1}^{q_2} \cdots \sum_{k_{(n_i+1)=1}}^{q_{(n_i+1)}} g_{k_1}^{(1)} g_{k_2}^{(2)} \cdots g_{k_{(n_i+1)}}^{(n_i+1)} \prod_{j=1}^{n_i} f(y_{ij}^*|x_{kj}, z_{ij}, \tau_k; \beta) f(w_{ij}|x_{kj}; \sigma_e^2)]. \quad (18)$$

Under the assumption of misspecifying the measurement error distribution, in order to evaluate the integral (17) which includes only random effects τ_i as an integrated variable, a q-length vector comprise of quadrature points and the associated weights should be calculated from a Hermite polynomial with degree q. So, in this case, the set of GHQ nodes which are replaced with the random effects τ_i, with the corresponding weights, are defined in the following way:

$$\{\tau_k = (\tau_1, \cdots, \tau_q)', g_k = (g_1, \cdots, g_q)'\}. \quad (19)$$

As a notation, τ_t $(t = 1, \cdots, q)$ is the t-th zero of the Hermite polynomial, and g_t is the quadrature weight related to the quadrature point τ_t. With the associated nodes and weights in (19), the GHQ approximation of the log-likelihood function (17) will be of the following form:

$$l(\theta; y, w, z) = \sum_{i=1}^{m} \log[\sum_{k=1}^{q} \prod_{j=1}^{n_i} f(y_{ij}|w_{ij}, z_{ij}, \tau_k; \beta) g_k]. \quad (20)$$

The approximation of log-likelihood function corresponding to (12) and (13) can be obtained in the same direction.

4 Simulations and Inferences

In this section, we are going to conduct simulation studies in order to compare and appraise the efficiency of the Multivariate (Multidimensional) Gauss-Hermite Quadrature (MGHQ) approximation method in the analysis of multilevel data with measurement error in the covariates and misclassification in the binary response, i.e. the hybrid method to three other scenarios in which, classification error in the binary response or measurement error in the covariates or both of them, have been misspecified. We aim to investigate how measurement error in the covariate in conjunction with classification error might result in bias, root of mean squared error (RMSE) and coverage probability/rate (CR). We are interested to model the relationship between binary responses and two continuous covariates. The response variables are generated from Bernoulli distribution. Following is the logistic random effects regression model, used for data generating

$$logit[P(Y_{ij} = 1|x_{ij}, z_{ij}, \tau_i)] = \beta_0 + \beta_x x_{ij} + \beta_z z_{ij} + \tau_i, \tag{21}$$

for $i = 1, \cdots, m$ individuals and $j = 1, \cdots, n_i$ occasions. Such techniques have been implemented in the paper, using R 3.5.3 (R Core Team (2019)). Further, we assume $(\beta_0, \beta_x, \beta_z)$ are the intercept and covariates's fixed effects coefficients. We assume that z_{ij} is an error-free or exactly measured covariate generated as a normally distributed variable following $N(0, 4^2)$ distribution, which is treated fixed during the simulation study. The measurement error model that we have considered is the classical additive model. In this case, we generate a surrogate variable W_{ij} for the error-prone covariate X_{ij}, as in (2), where e_{ij} is the measurement error variable independently and identically distributed following $N(0, \sigma_e^2)$, Albeit we will change the magnitude of σ_e^2 (measurement error in X_{ij}) along with different scopes of misclassification rates during the simulation study. This is because we are going to check the impact of considering measurement error and classification error on the estimation procedure. It is important to point out that only the results corresponding to $\sigma_e^2 = 0.5$ will be recounted. The true but latent covariate X_{ij} is assumed to have a structural modeling, generated from a parametric homogenous normal distribution, so that, $X_{ij} = \mu_x + V_{ij}$, where the V_{ij} variables are independent following $N(0, \sigma_{xj}^2)$ for $j = 1, 2, 3$, and we have set $\Sigma_x = I_3$, i.e. identity matrix with three elements corresponding to each occasion measurement. Moreover, we consider random effects τ_i to have homogenous variance, that is $\tau_i \sim N(0, \sigma_\tau^2)$. In this direction, we consider m=1000 subjects with $n_i = 3$, as the number of follow-up for each individual. The vector of initial values including the fixed effects and the random effects variance are set to the following values:

$$\theta = (\beta_0, \beta_x, \beta_z, \sigma_\tau^2)' = (0.02, 0.6, 0.5, 2.4)'.$$

After generating Y_{ij} from Bernoulli distribution with probability of success P_{ij}, at the moment in the existence of classification errors, it can be concluded that under non-differential misclassification structure,

$$P_{ij}^* = \gamma_{01} + (1 - \gamma_{10} - \gamma_{01})P_{ij}, \tag{22}$$

Recalling that γ_{01} and γ_{10} are the misclassification probabilities. Remarkably, after generating P_{ij} from (21), in order to cover misclassification probabilities, P_{ij}^* are then generated by different set up of γ_{01} and γ_{10} by applying (22).

Due to the fact that there are latent true covariate X_{ij} for each of three different occasions and random effects τ_i needed to be integrated out in order to approximate the likelihood function via the MGHQ method, the number of quadrature points used in each of the dimensions was set to 3, Note that this restriction leads to generate eighty one nodes for both X_{ij} and τ_i integrands, along with a vector of weights with length equal to eighty one. It is important to note that the likelihood function is approximated with the MGHQ approximation method and then Newton-Raphson algorithm can be applied to calculate the MLEs of both fixed and random effects. After numerically approximating the likelihood function, the score function and observed information matrix can be automatically calculated. Then, the variance components of the MGHQ estimates are derived by inverting the negative of the Hessian matrix.

In order to check how covariate measurement error and misclassification in the response might influence the estimation procedure, we generate R=100 different data sets, each simulated data set contains $m = 1000$ individuals with $n_i = 3$ replicate measurements at each occasion. The aforesaid independent data sets are generated for fixed values of w_{ij} and z_{ij}. We then fit four distinct scenarios. The first one defines a model ignoring the covariate measurement error and response classification errors (naive method) and using the observed values of X_i and Y_i. The second scenario is the method that merely incorporates misclassification rates and dismisses covariate measurement error in the likelihood function (misclassification adjusted method or MIC in abbreviation). In the third situation, measurement error in the covariate will be accounted, while omitting the classification errors of the response variable in the estimation procedure (measurement error adjusted method or ME in simplification). The last script examines our proposed methodology, that concurrently amalgamates the measurement error process in the covariates and misclassification in the binary response in estimating the parameters of interest (hybrid method), also considering homogenous variance for the random effects.

4.1 Simulation Results

In this connection, we express simulation results, and evaluate the performance of the four methods described former. The results are provided in terms of the mean of absolute bias,

$$Bias_{\theta_0} = \frac{1}{R} \sum_{r=1}^{R} |\hat{\theta_0}^{(r)} - \theta_0|,$$

where θ_0 is considered as an initial and true value of a particular parameter, and $\hat{\theta_0}^{(r)}$ is the estimated value of θ_0 in the r-th simulation run, empirical standard error (SE) or standard deviation of the estimates over the simulations and root of mean squared error, i.e.

$$RMSE = \sqrt{Bias_{\theta_0}^2 + Var_{\theta_0}},$$

where Var_{θ_0} is the average of $\mathrm{Var}(\hat{\theta}_0)$ over R simulation runs. Furthermore, it is important to clarify that the coverage probability is calculated as the proportion of times that the true value of a parameter is covered by the 95% nominal approximate confidence interval during the simulations.
In Table 1, the outcomes of the four stated methods (naive, hybrid, ME and MIC) for R=100 simulated data sets are provided based on small magnitudes of covariate measurement error and response misclassification rates (small error with $\sigma_e^2 = 0.5$ and $\gamma_{01} = \gamma_{10} = 0.05$). The simulation results for the fixed effects are characterized by $(\beta_0, \beta_x, \beta_z)$. Furthermore, the estimation of random effects variance, measurement error variation and misclassification

probabilities are indexed as σ^2_τ, σ^2_e, γ_{01} and γ_{10}, respectively.

From the results of Table 1, with small degrees of error in covariates and the response variable, it is entirely understandable that considering and modifying for measurement error and classification error results in less biases in estimating the fixed effects in comparison with the naive, ME and MIC approaches (0.0001 for β_0, 0.0128 for β_x, and 0.0131 for β_z). Besides, the hybrid approach shows smaller RMSEs and better CRs for the nominal value confidence interval in estimating the fixed effects in contrast with other three methods. In such a suggested method, the estimation of measurement error variance together with misclassification rates was found significant for the 95% confidence interval. It is important to notice that based on the results of the ME approach, it can be achieved that neglecting the classification error in the response leads to underestimating the regression parameters. Moreover, as stated in the MIC method, it is derived that excluding the covariate measurement error and adjustment for classification error leads to reduction in bias in parameter estimation compared with the ME approach.

Taking into accout small measurement error (σ^2_e=0.5) and moderate misclassification probabilities ($\gamma_{01} = \gamma_{10} = 0.085$), we can percieve considerable biases in estimating the fixed effects in using the naive, ME and MIC approaches. The outcome also delineates small biases, RMSEs and good CRs for the hybrid methodology in contrast with other approaches. The results based on the proposed method reveals that the estimated misclassification probabilities in addition to the measurement error variance is remarkable with the estimated reliability ratio, i.e. $\hat{\lambda_j} = \sigma^2_{xj}/(\sigma^2_{xj} + \hat{\sigma^2_e}) = 0.6059$ with sensitivity 0.9171 and specificity equal to 0.9198.

Compared with the results of Table 1, we can infer the fact that with the increase in the misclassification rates, the biases tend to increase in the hybrid approach. According to the results of the ME approach, it can be concluded that adjustment for measurement error underestimates the regression coefficients. Table 2 also hints the fact that based on the results of the MIC aspect, adjustment for response misclassification refines the parameter estimation process. By way of illustration, the MIC approach results $\hat{\beta}_x$=0.5573 as the estimated value for β_x with bias 0.0427, while the ME method yields $\hat{\beta}_x$=0.3655 with bias 0.2345 as well as the naive approach with $\hat{\beta}_x$=0.3713 and bias value equal to 0.2287.

According to the results of Table 3 with small covariate measurement error (σ^2_e=0.5) and severe (i.e. high) misclassification rates ($\gamma_{01} = \gamma_{10} = 0.1$), it is utterly obvious that adjustment for covariate classical error and mis-

Table 1. Parameter estimation (Est), absolute of bias (Bias), standard error (SE), root of mean squared error (RMSE), and coverage rate (CR) of the parameter estimates with small degree of misclassification rates, $n_i = 3$ replicates and 100 simulation runs in accordance with the mentioned approaches. Naive approach misspecifies both sources of errors, ME approach ignores response misclassification, MIC approach dismisses covariate measurement error and the hybrid method incorporates both errors in the analysis of multilevel data.

		Naive					**Hybrid**				
Scale of Error	Parameter	Est	Bias	SE	RMSE	CR	Est	Bias	SE	RMSE	CR
small error											
($\sigma_e^2 = 0.5$)	β_0	0.0112	0.0087	0.0557	0.0564	0.85	0.0201	0.0001	0.0557	0.0557	0.94
	β_x	0.4341	0.1659	0.0419	0.1711	0.92	0.5872	0.0128	0.0834	0.0844	0.96
	β_z	0.3694	0.1306	0.0197	0.1320	0.89	0.4869	0.0131	0.0483	0.0500	0.91
	σ_τ^2	1.3084	1.0916	0.2255	1.1146	0.90	2.2574	0.1425	0.4589	0.4805	0.91
	σ_e^2	-	-	-	-	-	0.6556	0.1557	0.0292	0.1584	0.90
	γ_{01}	-	-	-	-	-	0.0465	0.0034	0.0232	0.0234	0.96
	γ_{10}	-	-	-	-	-	0.0425	0.0075	0.0213	0.0226	0.93
AIC				-3198.778					-13023.24		

		ME					**MIC**				
Scale of Error	Parameter	Est	Bias	SE	RMSE	CR	Est	Bias	SE	RMSE	CR
small error											
($\sigma_e^2 = 0.5$)	β_0	0.0086	0.0113	0.0614	0.0625	0.91	-0.0079	0.0279	0.1499	0.1501	0.94
	β_x	0.4402	0.1597	0.0413	0.1649	0.93	0.5716	0.0284	0.0865	0.0911	0.94
	β_z	0.3650	0.1349	0.0214	0.1365	0.91	0.4791	0.0209	0.0603	0.0638	0.90
	σ_τ^2	1.3066	1.0933	0.2639	1.1246	0.90	2.1723	0.2277	0.6390	0.6783	0.91
	σ_e^2	0.6563	0.1563	0.0274	0.1586	0.90	-	-	-	-	-
	γ_{01}	-	-	-	-	-	0.0372	0.0127	0.0111	0.0389	0.96
	γ_{10}	-	-	-	-	-	0.0395	0.0105	0.0134	0.0417	0.96
AIC				-12958.18					-13010.296		

Table 2. Parameter estimation (Est), absolute of bias (Bias), standard error (SE), root of mean squared error (RMSE), and coverage rate (CR) of the parameter estimates with moderate degree of misclassification rates, $n_i = 3$ replicates and 100 simulation runs in accordance with the mentioned approaches. Naive approach misspecifies both sources of errors, ME approach ignores response misclassification, MIC approach dismisses covariate measurement error and the hybrid method incorporates both errors in the analysis of multilevel data.

		Naive					**Hybrid**				
Scale of Error	Parameter	Est	Bias	SE	RMSE	CR	Est	Bias	SE	RMSE	CR
small error											
($\sigma_e^2 = 0.5$)	β_0	0.0381	0.0181	0.0190	0.0262	0.93	0.0179	0.0020	0.0015	0.0025	0.95
	β_x	0.3713	0.2287	0.0389	0.2319	0.93	0.5913	0.0087	0.0747	0.0752	0.95
	β_z	0.3095	0.1905	0.0155	0.1911	0.90	0.4892	0.0108	0.0530	0.0541	0.95
	σ_τ^2	0.9398	1.4602	0.1840	1.4717	0.92	2.3096	0.0904	0.5040	0.5120	0.92
	σ_e^2	-	-	-	-	-	0.6513	0.1513	0.0339	0.1550	0.91
	γ_{01}	-	-	-	-	-	0.0802	0.0047	0.0258	0.0262	0.98
	γ_{10}	-	-	-	-	-	0.0829	0.0021	0.0247	0.0248	0.95
AIC				-3430.2268					-13202		

		ME					**MIC**				
Scale of Error	Parameter	Est	Bias	SE	RMSE	CR	Est	Bias	SE	RMSE	CR
small error											
($\sigma_e^2 = 0.5$)	β_0	0.0130	0.0070	0.0575	0.0580	0.93	0.0089	0.0111	0.1617	0.1620	0.93
	β_x	0.3655	0.2345	0.0373	0.2374	0.91	0.5573	0.0427	0.0999	0.1087	0.91
	β_z	0.3074	0.1926	0.0177	0.1934	0.91	0.4733	0.0266	0.0758	0.0803	0.94
	σ_τ^2	0.8988	1.5012	0.2063	1.5153	0.90	2.2425	0.1574	0.8198	0.8348	0.92
	σ_e^2	0.6568	0.1569	0.0337	0.1604	0.90	-	-	-	-	-
	γ_{01}	-	-	-	-	-	0.0740	0.0110	0.0287	0.0308	0.95
	γ_{10}	-	-	-	-	-	0.0699	0.0151	0.0294	0.0330	0.93
AIC				-13149.26					-13162.15		

classified response data results in less biases in estimating the fixed effects parameters in comparison with the naive, ME amd MIC methods (0.0005 for β_0, 0.0226 for β_x and 0.0147 for β_z). Besides, the hybrid approach represents smaller RMSEs and superior CRs for the 95% confidence interval in estimating the parameters of interest unlike other scenarios. In such a case, according to the Wald interval, the estimation of measurement error variance, sensitivity and specificity have been identified significant. As previously mentioned, the biases in fixed effects estimates are larger for the misspecification of classification errors methods, i.e. the naive and ME approaches.

According to the results of Table 1, Table 2 and Table 3, it can be deduced that the expansion of misclassification rates results in an augmentation of biases in the estimation process regarding to methods neglecting classification error in the response, i.e. the naive and ME approaches. Moreover, as stated by the outcomes of the ME approach, it can be established that without taking into account the misclassified responses, there is an underestimation in the regression coefficients, and with the enhancement of misclassification rates, there occurs more underestimation for the model parameters. In addition, the results suggest the certainty that the MIC approach improves parameter estimation as against the naive and the ME frameworks. In view of the fact that the measurement error variance is estimated from the hybrid and ME methodologies, it can be assessed that this parameter has been overestimated along with smaller CR in contrast with other parameters, which might be caused due to the number of submitted parameters in the corresponding model and the ability of the model to fit other regression parameters.

We have also increased the scale of covariate measurement error to σ_e^2=1.3 and σ_e^2=2, in order to check the performance of methodologies not adjusting for covariate measurement error, i.e. the naive and the MIC approaches. Based upon the simulation results, it can be determined that even with small degree of misclassification rates ($\gamma_{01} = \gamma_{10} = 0.05$), the achievements of naive and MIC approaches become more attenuated, as a result of not adjusting for covariate measurement error in the estimation procedure.
With the intention of detecting a model that diminishes the information loss, we have calculated the Akaike Information Criterion (AIC) value as a goodness of fit test evaluated by the logarithm of the likelihood function for the four reviewed models in the estimation technique. The mean of the AIC value over simulations illustrates that the proposed hybrid approach

Table 3. Parameter estimation (Est), absolute of bias (Bias), standard error (SE), root of mean squared error (RMSE), and coverage rate (CR) of the parameter estimates with severe degree of misclassification rates, $n_i = 3$ replicates and 100 simulation runs in accordance with the mentioned approaches. Naive approach misspecifies both sources of errors, ME approach ignores response misclassification, MIC approach dismisses covariate measurement error and the hybrid method incorporates both errors in the analysis of multilevel data.

		Naive					**Hybrid**				
Scale of Error	Parameter	Est	Bias	SE	RMSE	CR	Est	Bias	SE	RMSE	CR
small error											
($\sigma_e^2 = 0.5$)	β_0	0.0105	0.0094	0.0476	0.0485	0.95	0.0205	0.0005	0.0102	0.0102	0.97
	β_x	0.3485	0.2514	0.0396	0.2545	0.80	0.5773	0.0226	0.0843	0.0873	0.92
	β_z	0.2860	0.2139	0.0156	0.2144	0.88	0.4853	0.0147	0.0565	0.0584	0.94
	σ_τ^2	0.7890	1.6109	0.1877	1.6218	0.90	2.2652	0.1348	0.5508	0.5670	0.93
	σ_e^2	-	-	-	-	-	0.6544	0.1545	0.0268	0.1568	0.90
	γ_{01}	-	-	-	-	-	0.0921	0.0079	0.0284	0.0295	0.95
	γ_{10}	-	-	-	-	-	0.0983	0.0071	0.0226	0.0227	0.94
AIC				-3566.212					-13187.52		

		ME					**MIC**				
Scale of Error	Parameter	Est	Bias	SE	RMSE	CR	Est	Bias	SE	RMSE	CR
small error											
($\sigma_e^2 = 0.5$)	β_0	0.0131	0.0069	0.0508	0.0512	0.95	0.0473	0.0274	0.1667	0.1689	0.96
	β_x	0.3426	0.2574	0.0411	0.2606	0.91	0.5482	0.0518	0.1150	0.1261	0.92
	β_z	0.2828	0.2171	0.0146	0.2175	0.90	0.4734	0.0266	0.0965	0.1001	0.92
	σ_τ^2	0.7585	1.6415	0.1619	1.6494	0.90	2.2418	0.1582	0.9032	0.9169	0.93
	σ_e^2	0.6550	0.1550	0.0266	0.1572	0.90	-	-	-	-	-
	γ_{01}	-	-	-	-	-	0.0919	0.0081	0.0313	0.0323	0.95
	γ_{10}	-	-	-	-	-	0.0916	0.0084	0.0337	0.0347	0.94
AIC				-13105.53					-13160.914		

causes reduction in the information leakage and provides a better fit than the naive, ME and MIC methods for all three various levels of misclassification probabilities.

5 Application: Contraceptive Use Status in Bangladesh

Since 1971, Bangladesh has continually confirmed its mission to control population growth. This political mission is important in realizing the attainment and success of family planning program in Bangladesh (Mitra and Al-Sabir, 1996). As a result of using modern and effective methods of contraception, Bangladesh has encountered a magnificent decline in maternal mortality as well as child-bearing rates. According to Lapham and Mauldin (1984), family planning programmes contributed to the decline in maternal fertility, however, other experts believe that a family planning programme can not have a major impact on population reduction. According to Becker

(1991), an improvement in contraceptive or birth control methods is principally an induced response to other decreases in the desire for children. However, for the policy consequences purpose, it is fundamental to distinguish the factors associated with the use of contraception techniques.
It is important to consider the fact that due to the deficiency of data associated with the analysis carried out in the ensuing article, i.e. scarcity of available data corresponding to correlated binary response, in addition with validation data sources to overcome with the issue of measurement error in the covariates and also classification error in the binary response, the proposed method has been applied to analyze a dataset from the contraception status utilized in Bangladesh, as an illustration. Although the suggested methodology has been formulated to multilevel data structure, it can also be assigned to data with a longitudinal design together with family studies, in which the association of parameters is a subject of interest. In Table 4, we have cited information about some factors that might be incorporated in modelling contraceptive behavior in Bangladesh. According to the descriptive statistics of the data in Table 4, it can be concluded that 39.73% of women (or their husbands) were using one of the available contraceptive methods (mainly pills in Bangladesh) and 60.27% were not using any contraception.
Since multilevel models are well qualified of recognizing hierarchical variation among regions, following the idea in Hossain (2005) and Rasbash et.al (2009), we will examine the geographical district influences by fitting a logistic random effects model. The main covariates incorporated in this analysis are womens' age, region of residence (urban, and rural as the reference), religion (Hindu, and Muslim as the reference), having children and women's education. The main reason for taking these covariates in consideration is because they are found to have major effects in contraceptive behaviour among married women in some researches. Based on the results of observed data in Table 4, it has been concluded that the use of contraceptive prevalence involves the married women aged between 15-49 years who were currently using at least one method of contraception with mean 29.3 and standard deviation 8.7. Women's age plays an important role in using contraceptive methods. According to Hossain (2005), middle aged women has higher chance of using contraception and as they become older, the use of contraception methods decreases. The women in the analysis mostly have low level of education with 2.4 children. In the subsequent, we consider the situation where there has been inaccurate measurement occurrence in recording the age along with

Table 4. Descriptive statistics for the variables associated in the analysis, Bangladesh, 1989.

Variables	Frequency(%)	Mean(Sd)
Contraceptive status		
Using Contraception	1139 (39.73)	-
Not using contraception	1728 (60.27)	-
Age	-	29.3076 (8.6998)
No. children in the family	-	2.4405 (1.0528)
District		
Urban	804 (28.04)	-
Rural	2063 (71.96)	-
Religion		
Muslim	2480 (86.5)	-
Hindu	387 (13.49)	-
Education level		
Lower primary	1806 (62.99)	-
Upper primary	439 (15.31)	-
Secondary and above	265 (9.24)	-
None	357 (12.45)	-
Total	2867	

Abbreviation: Sd referes to standard deviation. No. is Number of.

misclassification in contraceptive status records for the women covered in the analysis, no matter of the region of residence.

In order to have a precise and rigorous analysis in the contraceptive behaviour dataset in Bangladesh, there are some important factors to be included in the model, which have been originated from the quality of family planning service in recording the covariates. According to DeGraff (1991), the development of family planning services is not the only reason in increasing contraceptive use, this matter is principally due to social and economic conditions in Bangladesh, which is predominantly rural and economically reliant on agriculture; so desired family size is so high and children are valuable in the family for their beneficial role in production. Due to these facts, infant and child mortality rates are relatively high; and education levels are very low, mainly for women. These factors are consistently related with the decrease in the quality of family planning services and policy. In this regard, we consider the fact that there has been inaccurate measurement episode in documenting the covariate age for the women envolved in the analysis. Hence, we focus on the perception where there exists an unwilling fault in age measurement reported by the interviewers or informed by the women associated in the interview. Based on the records of the family planning services, it can be realized that some community health workers have reported induced abortions as spontaneous abortions, i.e. using contraception methods to terminate pregnancy have been reported as not using contraception. As a consequence of this matter, the number of abortions in each area is underestimated (Johnston, 1999). Therefore, there seems to be some incorrect definition of each of possible outcomes of a pregnancy, including spontaneous abortions (not using contraception) and induced abortions (using contraception methods). This misclassification appears to be consistent over time and also between districts.

Let Y_{ij} be a binary response variable, with two categories comprising the contraceptive use status at the time of survey, i.e. whether the woman was using contraception or not using birth control methods, which has been decided as the reference category in the model. The random intercept model we have handled, is of the following form:

$$\begin{aligned} logit[P(Y_{ij}=1)] &= \beta_0 + \beta_1 Age_{ij} + \beta_2 Urban_{ij} + \beta_3 Hindu_{ij} \\ &\quad + \beta_4 Lc_{ij} + \beta_5 Educ_{ij} + \tau_j. \end{aligned} \tag{23}$$

The purpose of the analysis is to identify factors associated with use of contraception and also to inspect the scope of between-district variation in contraceptive use, concurrently accounting for measurement error in covariate and misclassification in the binary response. The data have a two-level hierarchical structure, with $i = 1, \cdots, 2867$ women (level 1) nested in $j = 1, \cdots, 60$ districts (level 2). It deserves to note that $Urban_{ij}$ is a binary covariate indicating the residence area of individual i at district j, as well as $Hindu_{ij}$ showing whether the individual's religion is Muslim or Hindu, Lc_{ij} is the number of living children in the family, and $Educ_{ij}$ is the education level of woman i at district j. In this setting, Age_{ij} represents the true age of individual i at district j, which is latent and treated as the error-prone covariate. The Kolmogorov-Smirnov test has been implemented to test the normality of variable Age. The value of the test statistic is $D = 0.16778$ with p-value=0.3295, accordingly accepting the normality assumption for the Age covariate. It is assumed that W_{ij} is the age recorded for woman i at district j, following classical additive structural measurement error model $W_{ij} = Age_{ij} + e_{ij}$. It is assumed that $Age_{ij} \sim N(\mu_x, \sigma_x^2)$, and the measurement error variable follows independent normal distribution with mean

0 and variance σ_e^2. In our illustration, for the misclassification process, we have assumed that Y_{ij}^* is the observed outcome and Y_{ij} is the true binary response variable of woman i at district j. In such a case, misclassification probabilities have been defined as in (4) and (5). In the present case, in order to specify the district effects on the probability of using contraception methods, it is assumed that the distribution of district effects τ_j is normal with mean 0 and variance σ_τ^2.

The proposed hybrid approach in this article has been employed for incorporating covariate measurement error along with response classification error in the estimation of regression coefficients and variance components. The results of parameter estimation and corresponding standard errors as well as the 95% approximated interval based upon the hybrid likelihood approach are displayed in Table 5.

From the results of Table 5, it can be concluded that age effect is significantly negative and supporting also by the 95% approximated interval, older women are less expected to use contraception. The women in urban areas, seem to be more probable to use contraception than women in rural areas, because of significant positive effect of the covariate Urban. Moreover, Table 5 indicates that there is a significant variation in terms of religious viewpoint. The Hindu effect is positive and significant, so it can be claimed that Hindu women are more expected to use contraception than Muslim women. This is mainly due to the fact that religious beliefs can decrease the propensity of contraceptive behaviour. In addition, Table 5 demonstrates that there is a significant variation between the increase of living children in the family and using contraception other than not using contraception. Hence, according to the results, women with children turn up significantly more expected to use contraception compared with childless ones. The education impact is positive and also significant according to the 95% approximated interval, announcing that more educated women are more probable to use contraception methods. Using the data in Table 5, it can be demonstrated that the intraclass correlation coefficient (ICC) statistic, in terms of fundamental latent response, is calculated as $\frac{\hat{\sigma}_\tau^2}{\hat{\sigma}_\tau^2+\pi^2/3} = 0.0956$, accordingly the expected correlation in the tendency to use contraception between two women from the same district is estimated 0.0956. This extent can be interpreted in a manner that about 9.6% of the variation in womens' inclination to use contraception lies between districts. For the Bangladesh data, the estimated reliability ratio is 0.6178, that is to say $\frac{\hat{\sigma}_x^2}{\hat{\sigma}_x^2+\hat{\sigma}_e^2} = 0.6601$, which expresses that there is about 33.98% error joined with the covariate age. In addition, we

Table 5. Coefficients, standard errors (SE), lower bound (LB) and upper bound (UB) of the 95% confidence interval based on the hybrid approach for the contraceptive status other than not using contraception, Bangladesh, 1989.

	Contraception Consumption			
Explanatory Variables	Coefficient	SE	LB	UB
Constant	-3.5324*	1.7174	-6.8985	-0.1663
Age	-0.0212*	0.0073	-0.0355	-0.0068
Residence				
Rural	-			
Urban	0.8866*	0.3733	0.1549	1.6183
Religion				
Muslim	-			
Hindu	0.6949*	0.3028	0.1014	1.2883
No. Children in the family	0.5878*	0.2285	0.1399	1.0356
Education level	0.4217*	0.1200	0.1865	0.6569
	Coefficient	SE	LB	UB
μ_x	29.3014*	0.1879	28.9331	29.6697
σ_x^2	50.1955*	2.0186	46.2390	54.1519
σ_τ^2	0.3478*	0.0304	0.2882	0.4074
σ_e^2	25.8411*	1.6549	22.5975	29.0847
γ_{01}	0.0784*	0.0280	0.0235	0.1332
γ_{10}	0.1730*	0.0856	0.0052	0.3407

Abbreviation: * refers to Significant at 0.05, - shows the reference level.

estimated that $P(Y_{ij}^* = 1|Y_{ij} = 1) = 0.827$ and $P(Y_{ij}^* = 0|Y_{ij} = 0) = 0.922$, reflecting the estimated sensitivity and specificity, respectively.

Conclusions

Multilevel studies are efficient statistical inference methods in the analysis of data that are organized in more than one level, as noted by Goldstein (2011). By applying the clustering information in the multilevel models, researchers can provide correct standard errors, confidence intervals and also significance tests, and these generally will be more ‘‘conservative’’ than the traditional models that are taken into account by simply ignoring the presence of hierarchy in the data structure. To make relevant statistical inference in the multilevel setting, many methods have been applied. The application of these methods relies on the fact that all variables in the study are correctly measured. Nonetheless, in many situations, this assumption does not hold for some of continuous covariates. Measurement error in any of the covari-

ates results in biased coefficients and incorrect inferences in estimating the parameters. An exclusive type of measurement error for discrete variables is called misclassification. Insensitivity of errors in either covariate or response variable is, not always plausible.
In this paper, we proposed the multivariate Gauss-Hermite quadrature approximation method for the likelihood inference of the random effects logistic regression model with repeatedly measured covariates subject to classical additive measurement error and non-differential misclassification assumption in the binary outcome. To highlight the impact of involving measurement error in the covariates in conjunction with classification error in the response variable in the analysis of multilevel data with binary response, we compared the suggested hybrid approach in this article to the methods where the covariate measurement error or response misclassification or both of these two errors have been misspecified.
Simulation studies indicated that ignoring measurement error and/or misclassification attenuates the estimates of regression coefficients. Parameter estimates appear to be sensitive with alteration of the degree of classification error probabilities, and the effect of misspecifying misclassification seems to be very destructive. Overally, according to the simultion results, the essential for reflection of an error adjusted methodology that corrects for classification error in the binary response together with measurement error in the covariates is apparent. Based on the real dataset analysis, we investigated the presence of measurement error in the covariate age, as well as misclassification in the contraceptive status response. To correct for these two sources of errors, we employed the proposed approach in this paper, using the multivariate Gauss-Hermite quadrature approximation technique.
Although the ensuing article generally assumed that the misclassification probabilities do not rely on covariates, i.e. non-differential misclassification process, there might be benefits of directly modelling the impact of associated covariates in the misclassification process. The approach presented here can be extended to cope with differentially misclassified response. Developing existing method to multilevel models containing measurement error in the covariate and differential classification error in the categorical response variable will be the subject of our forthcoming attemp.

Acknowledgment

The authors should appreciate the time and effort that reviewers as well as the respected editor-in-chief devoted to review our manuscript.

References

Agresti, A. (2002). *Categorical Data Analysis.* New York, John Wiley & Sons.

Becker, G. S. (1991). *A Treatise on the Family.* Cambridge, Harvard University Press.

Buonaccorsi, J.P. (2010). *Measurement Error, Models, Methods and Applications.* New York, CRC Press.

Buonaccorsi, J.P., Romeo, G., and Thoresen, M. (2018). Model-Based Bootstapping when Correcting for Measurement Error with Application to Logistic Regression. *Biometrics*, **74**, 135-144.

Carroll, R.J., Ruppert, D., Stefanski, L.A., and Crainiceanu, C.M. (2006). *Measurement Error in Nonlinear Models: A Modern Perspective.* Boca Raton, CRC Press.

Cheng, D., Branscum, A.J. and Stamey, J.D. (2010). Accounting for Response Misclassification and Covariate Measurement Error Improves Power and Reduces Bias in Epidemiologic Studies. *Annals of Epidemiology*, **20**, 562-567.

DeGraff, D. S. (1991). Increasing Contraceptive Use in Bangladesh: The Role of Demand and Supply Factors. *Demography*, **28**, 65-81.

Diggle, P.J., Liang, K.Y., and Zeger, S.L. (1994). *Analysis of Longitudinal Data.* Oxford, Oxford University Press.

Gerlach, R. and Stamey, J. (2007). Bayesian Model Selection for Logistic Regression with Misclassified Outcomes. *Statistical Modelling*, **7(3)**, 255-273.

Goldstein, H. (2011). *Multilevel Statistical Models (4th ed.).* London, John Wiley & Sons.

Hossain, M.B., (2005). Analysing the Relationship between Family Planning Workers' Contact and Contraceptive Switching in Rural Bangladesh Using Multilevel Modelling. *Journal of Biosocial Science*, **37**, 529-544.

Huq, N.M. and Cleland, J. (1990). *Bangladesh Fertility Survey*, National Institute of Population Research and Training (NIPORT).

Jaeckel, P. (2005). A Note on Multivariate Gauss-Hermite Quadrature. London: ABN-Amro. Retrieved from
http://www.pjaeckel.webspace.virginmedia.com/ANoteOnMultivariateGaussHermiteQuadrature.pdf

Johnston, H. (1999). *Measuring Induced Abortion: Integrating Qualitative with Quantitative Research Techniques.* Ph.D. thesis: The Johns Hopkins University.

Lapham, R. J. and Mauldin, W. P. (1984). Family Planning Effort and Birthrate Decline in Developing Countries, *International Family Planning Perspectives.* **10**, 109-l18.

Lele, S.R., Dennis, B., and Lutscher, F. (2007). Data Cloning: Easy Maximum Likelihood Estimation for Complex Ecological Models Using Bayesian Marcov Chain Monte Carlo Methods. *Ecology Letters*. **10**, 551-563.

Lyles, R.H., Tang, L., Superak, H.M., King, C.C., Celentano, D.D., Lo, Y. and Sobel, J.D. (2011). Validation Data-Based Adjustments for Outcome Misclassification in Logistic Regression: an Illustration. *Epidemiology (Cambridge, Mass.)*, **22**, 589.

Magder, L.S. and Hughes, J.P. (1997). Logistic Regression when the Outcome Is Measured with Uncertainty. *American Journal of Epidemiology*, **146**, 195-203.

McCullach, C.E. (1994). Maximum Likelihood Variance Components Estimation for Binary Data. *Journal of the American Statistical Association*, **89**, 330-335.

McCullach, C.E., Searle, S.R., and Neuhaus, J.M. (2008). *Generalized, Linear, and Mixed Models (2nd ed.)*. London, John Wiley & Sons.

Mitra, S.N. and Al-Sabir, A. (1996). DHS Working Papers.

Molenberghs, G., and Verbeke, G. (2006). *Models for Discrete Longitudinal Data*. New York, Springer Science & Business Media.

Nelder, J. A. and Mead, R. (1965). A Simplex Algorithm for Function Minimization. *Computer Journal*, **7**, 308- 313.

Neuhaus, J.M. (1999). Bias and Efficiency Loss Due to Misclassified Responses in Binary Regression. *Biometrika*, **86(4)**, 843-855.

Neuhaus, J.M. (2002). Analysis of Clustered and Longitudinal Binary Data Subject to Response Misclassification. *Biometrics*, **58(3)**, 675-683.

Pan, J. and Thompson, R. (2003). Gauss-Hermite Quadrature Approximation Estimation in Generalized Linear Mixed Models. *Computational Statistics*, **18**, 57-78.

Paulino, C. D., Soares, P. and Neuhaus, J. (2003). Binomial Regression with Misclassification. *Biometrics*, **59**, 670-675.

R Core Team. R (2019). A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

Rasbash, J., Steele, F., Browne, W. J., & Goldstein, H. (2009). *A User's Guide to MLwiN, v2. 10.* London: Institute of Education.

Roy, S. (2012). Accounting for Response Misclassification and Covariate Measurement Error Using a Random Effect Logit Model. *Communications in Statistics-Simulation and Computation*, **41**, 1623-1636.

Skrondal, A. and Rabe-Hesketh, S. (2004). *Generalized Latent Variable Modeling*. Boca Raton, CRC Press.

Stroup, W.W. (2012). *Generalized Linear Mixed Models: Modern Concepts, Methods and Applications*. Boca Raton, CRC Press.

Tang, L., Lyles, R.H., King, C.C., Hogan, J.W. and Lo, Y. (2015). Regression Analysis for Differentially Misclassified Correlated Binary Outcomes. *Journal of the Royal Statistical Society. Series C, Applied Statistics*, **64**, 433.

Tanner, M.A. (1993). *Tools for Statistical inference: Observed Data and Data Augmentation(2nd ed.)*. New York, Springer Science & Business Media.

Torabi, M. (2013). Likelihood Inference in Generalized Linear Mixed Measurement Error Models. *Computational Statistics and Data Analysis*, **57**, 549-557.

Wang, N., Lin, X., and Guttierrez, R.G. (1999). A Bias Correction Regression Calibration Approach in Generalized Linear Mixed Measurement Error Models. *Communications in Statistics*, **28**, 217-232.

Wang, N., Lin, X., Guttierrez, R.G., and Carroll, R.J. (1998). Bias Analysis and SIMEX Approach in Generalized Linear Mixed Measurement Error Models, *Journal of American Statistical Association*, **93**, 249-261.

Wu, L. (2009). *Mixed Effects Models for Complex Data*. Boca Raton, CRC Press.

Xie, X., Xue, X. and Strickler, H.D. (2017). Generalized Linear Mixed Model for Binary Outcomes when Covariates are Subject to Measurement Errors and Detection Limits. *Statistics in Medicine*, **37**, 119-136.

Yi, G.Y. (2016). *Statistical Analysis with Measurement Error or Misclassification*. New York, Springer.

Mousa Golalizadeh
Department of Statistics,
Tarbiat Modares University,
Tehran, Iran.
email: *golalizadeh@modares.ac.ir*

Maryam Ahangari
Department of Statistics,
Tarbiat Modares University,
Tehran, Iran.
email: *m.ahangari@modares.ac.ir*

Zahra Rezaei Ghahroodi
School of Mathematics, Statistics and Computer Science,
University of Tehran,
Tehran, Iran.
email: *z.rezaeigh@ut.ac.ir*