Published in final edited form as:

Biom J. 2017 January ; 59(1): 94–109. doi:10.1002/bimj.201600043.

A zero-augmented generalized gamma regression calibration to adjust for covariate measurement error: A case of an episodically consumed dietary intake

George O. Agogo*

Department of Internal Medicine, Yale University, 300 George St, Suite 775, New Haven, CT 06511, USA

Abstract

Measurement error in exposure variables is a serious impediment in epidemiological studies that relate exposures to health outcomes. In nutritional studies, interest could be in the association between long-term dietary intake and disease occurrence. Long-term intake is usually assessed with food frequency questionnaire (FFQ), which is prone to recall bias. Measurement error in FFQ-reported intakes leads to bias in parameter estimate that quantifies the association. To adjust for bias in the association, a calibration study is required to obtain unbiased intake measurements using a short-term instrument such as 24-hour recall (24HR). The 24HR intakes are used as response in regression calibration to adjust for bias in the association. For foods not consumed daily, 24HR-reported intakes are usually characterized by excess zeroes, right skewness, and heteroscedasticity posing serious challenge in regression calibration modeling. We proposed a zero-augmented calibration model to adjust for measurement error in reported intake, while handling excess zeroes, skewness, and heteroscedasticity simultaneously without transforming 24HR intake values. We compared the proposed calibration method with the standard method and with methods that ignore measurement error by estimating long-term intake with 24HR and FFQ-reported intakes. The comparison was done in real and simulated datasets. With the 24HR, the mean increase in mercury level per ounce fish intake was about 0.4; with the FFQ intake, the increase was about 1.2. With both calibration methods, the mean increase was about 2.0. Similar trend was observed in the simulation study. In conclusion, the proposed calibration method performs at least as good as the standard method.

Keywords

Attenuation; Excess zeroes; Generalized gamma model; Measurement error; Regression calibration

*Corresponding author: agogogeorge@gmail.com, Phone: +1-203-737-2908.

Conflict of interest

The author has declared no conflict of interest.

Additional supporting information including source code to reproduce the results may be found in the online version of this article at the publisher's web-site

1 Introduction

Measurement error in exposure variables is a major impediment in epidemiological studies that relate exposures to health outcomes such as in nutritional epidemiology (Kaaks et al., 2002). In dietary assessment, a food frequency questionnaire (hereafter, FFQ) is usually used to assess an individual's long-term average intake (hereafter, usual intake) (Kaaks and Riboli, 1997; Kaaks et al., 2002). The problem is that in the FFQ, past intake is recalled, usually within the past one year, leading to recall bias due to memory failure (Willett, 1998; Carroll et al., 2012). Due to measurement error in the FFQ intake, the estimate for the association between dietary intake and a health outcome will be biased and there will be loss in statistical power to detect a significant association (Kipnis et al., 1999, 2003; Freedman et al., 2011; Carroll et al., 2012).

Regression calibration is often used to adjust for the bias in the association caused by measurement error in the FFQ. Regression calibration involves computation of conditional expectation of unobserved true intake given the observed data (see Section 4 of the main text). In practice, a calibration study with unbiased measurements of true intake is required to estimate the calibration function as explained in Keogh and White (2014). The unbiased measurements that are used as the response values in the calibration model can be measured with short-term instruments such as 24-hour recall (hereafter, 24HR) (Slimani et al., 2000, 2002). In the 24HR, intake in the previous day prior to the interview time is recorded. For foods not consumed daily (hereafter, episodically consumed), many zeroes are recorded in the 24HR to denote nonconsumption. The distribution of such an episodically consumed food is usually characterized by excess zeroes, right skewness, and with heteroscedastic variance (Tooze et al., 2006; Kipnis et al., 2009; Rodrigues-Motta et al., 2015). Due to excess zeroes and large within-individual variability, using 24HR intake to relate episodically consumed food with a health outcome can be problematic (Kipnis et al., 2009). Likewise, using the FFQ intake alone can be misleading, due to sizable recall bias in the FFQ (Willett, 1998). The optimal strategy, therefore, is to combine intake data reported in the 24HR and FFQ together with individual characteristic variables to predict usual intake. The usual intake can be predicted with regression calibration method as described in Kipnis et al. (2009).

Regression calibration involves computation of conditional expectation of true intake given intake measured with error and a vector of error-free covariates (Carroll et al., 2012; Agogo et al., 2014). The estimated expected values are used, in place of the FFQ intake, to relate usual intake to a health outcome. As true usual intake is unobservable, the 24HR intake is used as the response variable in the calibration model. The challenge is how to handle excess zeroes, right skewness, and heteroscedasticity in the response variable. To handle excess zeroes, a two-part model is commonly used, where in part I, the mean probability of a nonzero response (hereafter, consumption probability) is modeled, and in part II, the mean of actual nonzero values (hereafter, consumed amount) is modeled (Olsen and Schafer, 2001; Yau et al., 2002; Tooze et al., 2006; Kipnis et al., 2009; Liu et al., 2010; Agogo et al., 2014, 2016). In part I, logistic or probit regression can be used. In part II, the standard method is to transform the nonzero response values to normality and hopefully, homoscedasticity, apply a linear (mixed) model on the transformed scale, predict intake on

the transformed scale, and transform the predicted intake back to the original scale (Tooze et al., 2006; Kipnis et al., 2009; Freedman et al., 2011). Transforming back to the original scale, however, can be problematic, especially in the presence of large within-person and between-person variations (Souverein et al., 2011). Recently, Rodrigues-Motta et al. (2015) proposed a zero-augmented mixture model to handle excess zeroes and skewness simultaneously without transforming the response. The authors, however, did not extend their model to adjust for exposure measurement error.

We proposed a flexible zero-augmented calibration model to correct for measurement error in intake reported in the FFQ and to handle excess zeroes, skewness, and heteroscedasticity simultaneously without transforming the 24HR response values. In the proposed model, heteroscedasticity is accounted for by modeling variance as a function of a set of individual-specific covariates.

Currently, to our knowledge, zero-augmented generalized gamma calibration model proposed here has not been extended to adjust for exposure measurement error. The proposed calibration method was compared with the standard method commonly used in the literature (Tooze et al., 2006; Kipnis et al., 2009), and with the methods that ignore measurement error. The methods were applied to NHANES 2003–2004 dataset (one of the largest nutritional studies in the United States) and in a simulation study.

The other sections of the paper are organized as follows. In Section 2, NHANES dataset is described. In Section 3, a generalized gamma distribution is described. In Section 4, regression calibration is described in a general context and in the context of excess zeroes, skewness, and heteroscedasticity. In Section 5, the methods that ignore measurement error are described. In Section 6, simulation study is described. In Section 7, the results are presented and discussed in Section 8.

2 NHANES data

The National Health and Nutrition Examination Survey (NHANES) is a collection of studies aimed to assess the effect of nutrition on health outcomes in the United States (CDC, 2015). For this study, we used data for women who participated in NHANES between 2003 and 2004. The ultimate goal was to investigate the association between usual fish intake and the level of mercury in the blood (similar to Kipnis et al., 2009). Data consisted of 1605 women aged between 12 and 49 years who provided one FFQ intake data and at least one 24HR intake data of fish and a blood sample for the measurement of serum mercury. In the study, 14.95% ($n = 240$) participants reported fish consumption in the first 24HR and 13.96% ($n = 216$) in the second 24HR. Moreover, 75.14% ($n = 1206$) reported no fish consumption in either 24HRs, whereas only 3.55% ($n = 57$) reported fish consumption in both 24HRs. Unlike in the 24HRs, almost everyone reported fish consumption in the FFQ with a percentage consumption of 96.76% ($n = 1553$). A spike at zero shown in Fig. 1 for the distribution of fish intake reported in the two 24HRs indicates that fish is not consumed daily by everyone.

In the preliminary analysis, weak correlation was observed between replicates of measurements from the 24HR and between intake measurements from 24HR and the FFQ (i.e., ρ is between 0.04 and 0.1). Individual characteristic variables considered include level of education and age of the participants. In Table 1, summary measures for the distribution of fish intake reported in the two 24HRs and FFQ are presented. Notably, higher percentage of zero intakes is reported in the 24HR than in the FFQ. This is due to a shorter recall period for the 24HR than for the FFQ. Additionally, the variability (standard deviation) of intake reported in 24HR is greater than the variability of intake reported in the FFQ. Lastly, a substantial degree of skewness is shown in the distribution of 24HR intake. This exploratory data analysis implies that using simple models, such as linear regression calibration, would be inappropriate in modeling the distribution of usual intake from the 24HR. In the next section, we described a generalized gamma distribution that is capable of handling skewness and heteroscedasticity complexities in the fish intake amount.

3 A generalized gamma distribution for a skewed and heteroscedastic variable

The data exploration for the distribution of 24HR intake suggested presence of excess zeroes and right skewness. We applied a generalized gamma distribution to handle right skewness and heteroscedasticity in the intake amount. The probability density function for the generalized gamma distributed random variable is defined by three parameters, σ , κ , and μ . Following the same parametrization used in Liu et al. (2010), Manning and Mullahy (2001), and Manning et al. (2005), the probability density function of a generalized gamma variable is parametrized as

$$f(R; \kappa, \mu, \sigma) = \frac{\gamma^\gamma}{\sigma R \Gamma(\gamma) \sqrt{\gamma}} \exp \left[\omega \sqrt{\gamma} - \gamma \exp(|\kappa| \omega) \right]$$

for $R > 0$, where R is intake amount from the 24HR, $\gamma = |\kappa|^{-2} > 0$, $\sigma > 0$, and $\omega = \text{sign}(\kappa) \{\log(R) - \mu\} / \sigma$, $\log(\cdot)$ is a natural logarithm function; $\Gamma(\cdot)$ is a gamma function. The parameters σ and κ quantify the degree of skewness (Manning et al., 2005). To allow for heteroscedasticity, the parameter σ is parameterized as a function of individual-specific covariates as described in Liu et al. (2010). In addition, intake amount can be influenced by these individual-specific characteristic variables. The parameter μ is modeled as a function of these characteristic variables to incorporate their influence. The expected value of the generalized gamma distributed variable R is given by

$$E(R) = \exp \left\{ \mu + \frac{\sigma \log(\kappa^2)}{\kappa} + \log \left[\Gamma \left(\frac{1}{\kappa^2} + \frac{\sigma}{\kappa} \right) \right] - \log \left[\Gamma \left(\frac{1}{\kappa^2} \right) \right] \right\}, \quad (1)$$

and variance by

$$\text{var}(R) = \left\{ \exp(\mu) \kappa^{2\sigma/\kappa} \right\}^2 \left\{ \frac{\Gamma(1/\kappa^2 + \frac{2\sigma}{\kappa})}{\Gamma(1/\kappa^2)} - \left[\frac{\Gamma(1/\kappa^2 + \frac{2\sigma}{\kappa})}{\Gamma(1/\kappa^2)} \right]^{-2} \right\}.$$

Further advantage of the generalized gamma distribution is that it accommodates some well-known standard distributions as special cases. The distributions that comprise special cases of the generalized gamma include standard gamma distribution ($\kappa = \sigma > 0$), Weibull ($\sigma > 0$ and $\kappa = 1$), exponential ($\kappa = \sigma = 1$), and log-normal ($\sigma > 0$ and $\kappa \rightarrow 0$) as explained in Manning et al. (2005). Noteworthy, besides the parametrization used here for the generalized gamma distribution, other possible parameterizations can be found in Stacy (1962), and Stacy and Mihram (1965).

4 Regression calibration

In nutritional epidemiology, the interest could be to investigate the association between long-term dietary intake and risk of a disease (hereafter, diet–disease association). The diet–disease association can be described by a generalized linear model given by

$$\vartheta\{E(Y_i|T_i, \mathbf{Z}_i)\} = \beta_T T_i + \beta_Z^T \mathbf{Z}_i, \quad (2)$$

where Y_i is a disease outcome for the i -th individual, for example, cancer, T_i is true usual dietary intake, $\mathbf{Z}_i = (Z_1, \dots, Z_k)^T$ is a vector of k error-free confounding variables with effect parameters $\beta_Z^T = (\beta_1, \dots, \beta_k)$, ϑ is a function that links the mean outcome with the linear predictor. In such nutritional study, the main interest is to estimate the parameter that associates usual intake to disease outcome, denoted by β_T in model (2). In practice, however, T_i is not observable; therefore, the FFQ intake (hereafter, denoted by Q_i) is usually used, leading to bias in estimating β_T .

To adjust for the bias in β_T using regression calibration, the FFQ intake is replaced with the conditional expectation of true intake given the FFQ intake and error-free covariates. True intake, however, is unobservable and is replaced with intake measurements from the 24HR in the calibration model. We denote the j -th replicate of intake measurement from the 24HR for the i -th subject by R_{ij} , $i = 1, \dots, n$; $j = 1, 2$. An important assumption in regression calibration is that R_{ij} is unbiased for T_i and is measured with error that is uncorrelated with measurement error in Q_i and with the true intake T_i . Regression calibration involves finding the best conditional expectation of true intake given Q_i and a set of covariates assumed to be measured without error as: $E(T_{ij}|Q_i, \mathbf{Z}_i^T, \mathbf{C}_i^T)$, where $\mathbf{C}_i = (C_{k+1}, \dots, C_p)^T$ is a vector of error-free covariates that conditionally predict T_{ij} given $\mathbf{Z}_i$ but do not predict Y_i . Because T_{ij} is unobservable, R_{ij} is used such that

$$T_i = E(T_{ij}|Q_i, \mathbf{Z}_i^T, \mathbf{C}_i^T) = E(R_{ij}|Q_i, \mathbf{Z}_i^T, \mathbf{C}_i^T). \quad (3)$$

In model (3), a nondifferential measurement error is assumed in Q_i with respect to Y_i that is, $f(Y_i|T_i, Q_i, \mathbf{Z}_i^T) = f(Y_i|T_i, \mathbf{Z}_i^T)$, where $f(\cdot)$ is a density function. This means that Q_i does not contain extra information about Y_i over what is contained in T_i (Carroll et al., 2006).

Fish is not consumed daily by everyone, leading to R_{ij} values with substantial amount of zeroes. Due to the excess zero R_{ij} values, model (3) can be reparametrized in two parts. In the first part, the mean probability of a nonzero R_{ij} is modeled as $P(R_{ij} > 0|Q_i, \mathbf{Z}_i^T, \mathbf{C}_i^T) = \pi_i$ (hereafter, part I). In the second part, the conditional mean of actual nonzero R_{ij} values is modeled as $E(R_{ij}|Q_i, \mathbf{Z}_i^T, \mathbf{C}_i^T; R_{ij} > 0) = A_i$ (hereafter, part II). Thus, the right-hand side of expression (3) can be reparametrized as the product of the two components:

$$E(R_{ij}|Q_i, \mathbf{Z}_i^T, \mathbf{C}_i^T) = P(R_{ij} > 0|Q_i, \mathbf{Z}_i^T, \mathbf{C}_i^T) \times E(R_{ij}|Q_i, \mathbf{Z}_i^T, \mathbf{C}_i^T; R_{ij} > 0) = \pi_i A_i. \quad (4)$$

In this paper, a logistic mixed model was applied in part I and a flexible generalized gamma mixed model in part II as described below.

4.1 A zero-augmented generalized gamma mixed regression calibration (proposed method)

Regression calibration model shown in expression (4) is specified on a vector of individual-specific covariates defined by $\mathbf{X}_i = (Q_i, \mathbf{Z}_i^T, \mathbf{C}_i^T)^T$ and correlated random effects u_{1i} in part I, and u_{2i} in part II. The random effects allow for an individual's probability of intake and consumed amount to deviate from the population profiles defined by the covariates $\mathbf{X}_i$.

In part I, the mean probability of a nonzero R_{ij} is modeled using a logistic mixed effects model as

$$\pi_i = \exp(\boldsymbol{\theta}^T \mathbf{X}_i + u_{1i}) / (1 + \exp(\boldsymbol{\theta}^T \mathbf{X}_i + u_{1i})), \quad (5)$$

where $\boldsymbol{\theta}$ is a vector of fixed effects parameters of the covariates and u_{1i} is independent of the covariates $\mathbf{X}_i$ and allows an individual's probability of consumption to deviate from the population mean profile defined by $\mathbf{X}_i$ (Kipnis et al., 2009).

In part II, a generalized gamma mixed effects model is used to estimate the conditional mean of actual nonzero R_{ij} values as described in Section 3. Because the mean of nonzero R_{ij} in part II is estimated conditional on individual covariates and random effects, the parameter μ in expression (1) is modeled as

$$\mu_i = \boldsymbol{\psi}^T \mathbf{X}_i + u_{2i}, \quad (6)$$

where $\boldsymbol{\psi}$ is a vector of parameters for the fixed effects of the covariates; u_{2i} is independent of the covariates $\mathbf{X}_i$ and allows for an individual's consumed amount to deviate from the population mean profile. Similarly, to allow for heteroscedasticity in the intake amount, the parameter σ in expression (1) is modeled as a function of individual-characteristic variables, $\mathbf{W}_i = (W_1, \dots, W_h)^T$ with a distribution function that guarantees only positive values, because variance is positive. Thus, σ is modeled with an exponential function as

$$\sigma_i = 0.5 \exp \left(\boldsymbol{\Delta}^T \mathbf{W}_i \right), \quad (7)$$

where $\boldsymbol{\Delta}$ is a vector of fixed effects of the covariates. Therefore, the conditional mean consumption amount A_i is obtained by using expressions (6) and (7) in (1). Subsequently, true usual intake for a specific individual is defined as

$$\pi_i A_i = \left[\frac{\exp \left(\boldsymbol{\theta}^T \mathbf{X}_i + u_{1i} \right)}{1 + \exp \left(\boldsymbol{\theta}^T \mathbf{X}_i + u_{1i} \right)} \right] \times \exp \left\{ \mu_i + \frac{\sigma_i \log(\kappa^2)}{\kappa} + \log \left[\Gamma \left(\frac{1}{\kappa^2} + \frac{\sigma_i}{\kappa} \right) \right] + \log \left[\Gamma \left(\frac{1}{\kappa^2} \right) \right] \right\}. \quad (8)$$

We, hereafter, refer to this model as the proposed model.

Importantly, the two parts of the proposed model are linked in two ways. First, they are linked by correlated between-person random effects $\mathbf{u}_i = (u_{1i}, u_{2i})^T \sim N(\mathbf{0}, \mathbf{D})$, where $\mathbf{D}$ is a positive definite covariance matrix and is given by

$$\mathbf{D} = \begin{pmatrix} \sigma_{u_1}^2 & \rho_{u_1 u_2} \sigma_{u_1} \sigma_{u_2} \\ \rho_{u_1 u_2} \sigma_{u_1} \sigma_{u_2} & \sigma_{u_2}^2 \end{pmatrix}.$$

The diagonal elements of $\mathbf{D}$ are the variance components and the off-diagonal elements are the covariance components. Second, the two parts are linked by overlapping covariates $\mathbf{X}_i$ on both parts of the model. Noteworthy, the two parts do not necessarily need to have the same set of covariates, but should share some covariates in order to induce the correlation between the two parts (hereafter, cross-part correlation). The cross-part correlation is common for episodically consumed foods; for instance, it is likely that an individual who consumes a given food more frequently might tend to consume more of that food on consumption events, resulting in a positive cross-part correlation.

4.2 Model fitting and assessment of goodness of fit

The proposed model was fitted by maximizing the likelihood of the data. To define the likelihood function, we let $\xi = (\theta^T, \psi^T, \kappa, \sigma_{u1}, \sigma_{u2}, \rho_{u1u2})^T$ be a vector of all parameters in the proposed model. Thus, these parameters are estimated by maximizing the following marginal likelihood function of the data:

$$\mathcal{L}(\xi) = \int \prod_{i=1}^n \prod_{j=1}^{m_i} \left[1 - P(R_{ij} > 0 | \mathbf{X}_i, u_{1i}; \theta^T) \right]^{I(R_{ij}=0)} \left[P(R_{ij} > 0 | \mathbf{X}_i, u_{1i}; \theta^T) \cdot f(R_{ij} | \mathbf{X}_i, u_{2i}; \Omega^T) \right]^{I(R_{ij}>0)} f(\mathbf{u}_i) d\mathbf{u}_i, \quad (9)$$

where $\Omega = (\psi^T, \kappa, \sigma_{u1}, \sigma_{u2}, \rho_{u1u2})^T$, $f(R_{ij} | \mathbf{X}_i, u_{2i}; \Omega^T)$ denotes the generalized gamma probability density function and $f(\mathbf{u}_i)$ denotes a probability density function for the random effects $\mathbf{u}_i$.

The proposed model was fitted using NLMIXED procedure in SAS version 9.4. The fitting was done by maximizing the above marginal likelihood of the data, which involves computation of the integral over the random effects. To carry out the computation, nonadaptive Gauss–Hermite quadrature method with 10 quadrature points was used and the derivative of the log-likelihood optimized using Newton–Raphson optimization technique. Starting values for the model parameters were obtained by fitting the two parts of the model separately using GLIMMIX procedure for part I and NLMIXED procedure for part II. Age and fish intake reported in the FFQ were included as the covariates on both parts of the proposed model.

To account for the uncertainty in the calibration process, the standard error for the parameter estimate that quantifies the association between intake and health outcome, β_T , was estimated using bootstrap method (Efron and Tibshirani, 1993). The bootstrap standard error is defined as the standard deviation of the association parameter estimates from the bootstrapped samples:

$$SE(\hat{\beta}_T) = \left\{ \frac{1}{B-1} \sum_{b=1}^B \left(\hat{\beta}_{T,b}^* - \overline{\hat{\beta}_T^*} \right)^2 \right\}^{0.5},$$

where $\hat{\beta}_T$ is the estimate of the association parameter, $\hat{\beta}_{T,b}^*$ is the parameter estimate from b -th bootstrap sample, and $\overline{\hat{\beta}_T^*}$ is the bootstrap mean estimate. We drew 1000 bootstrap samples. Each bootstrap sample of size 1000 was drawn from the original NHANES dataset without replacement, fitted the proposed calibration model and used the calibrated intake to estimate the association between mercury level in the blood and usual fish intake.

The goodness of fit of the calibration model was assessed using Akaike information criterion (AIC), defined as $AIC = 2p - \ln(L)$, where p is the number of parameters to be estimated and L is the maximized value of the likelihood function (Akaike, 1981). The reason for

using AIC is to be able to compare the proposed calibration model with the standard model commonly used in the literature. Note, these two calibration models are not nested within each other.

4.3 A standard calibration model with Box–Cox transformation (standard method)

Here, the standard calibration model commonly used in the literature is described (Tooze et al., 2006; Kipnis et al., 2009). In part I, a logistic mixed model is used. In part II, the actual nonzero R_{ij} values are Box–Cox transformed to normality and the conditional mean intake amount predicted on the transformed scale. The predicted intake values are transformed back to the original scale using the inverse of the Box–Cox transformation. With this model, true usual intake T_i is obtained as

$$\pi_i A_i = P(R_{ij} > 0 | \mathbf{X}_i, u_{1i}^*) \times E \left[g^{-1} \left(\psi^{*T} \mathbf{X}_i + u_{2i}^* + \varepsilon_{ij}; \lambda_R \right) | \mathbf{X}_i, u_{2i}^*; R_{ij} > 0 \right], \quad (10)$$

where ψ^* is a vector of fixed effects, u_{1i}^* and u_{2i}^* are vectors of random effects; $\varepsilon_{ij} \sim N(0, \sigma_\varepsilon^2)$ is normally distributed within-person random error that is independent of $\mathbf{X}_i$; $g^{-1}(X_i; \lambda_R)$ is the inverse of Box–Cox transformation and can be approximated by Taylor series expansion.

The Box–Cox transformation is defined as $g(X_i; \lambda_R) = (X_i^{\lambda_R} - 1)/\lambda_R$, if $\lambda_R \neq 0$, or $g(X_i; \lambda_R) = \log(X_i)$ if $\lambda_R = 0$. The data and SAS macro used to fit this model can be found in Supporting Information of Kipnis et al. (2009). This model is, hereafter, referred to as the standard model.

5 Methods that ignore measurement error (naïve methods)

We further estimated the association parameter by ignoring measurement error in the self-reported intake. In estimating the association parameter with the naïve methods, usual intake was estimated by (i) the average intake reported on the 24HRs (hereafter, mean-24HR method) and (ii) intake reported on the FFQ (hereafter, FFQ method). Noteworthy, with the naïve methods, we do not perform regression calibration.

6 Simulation study

The performance of the proposed model was assessed in a simulation study. To simulate true usual fish intake, two approaches were used. The difference between the two simulation approaches was on how the nonzero R_{ij} values were generated (consumed amount). The motivation for using different simulation approaches for the consumed amount was to assess how the proposed model performs when true consumed amount follows a different distribution from the generalized gamma distribution. The idea was to mimic the two distributions assumed for part II in the proposed and the standard models described in Sections 4.1 and 4.3, respectively. The simulation was conducted as follows.

First, the original NHANES data was bootstrapped to obtain 2000 bootstrap samples for the covariate matrix that consisted of age, level of education, and fish intake reported in the FFQ.

Second, normally distributed between-person random intake components for the probability of consumption (u_{1i}) and the consumed amount (u_{2i}) were generated using the covariance matrix estimate from NHANES data as

$$(u_{1i}, u_{2i})^T \sim N \left((0, 0)^T, D = \begin{bmatrix} 1.0296 & 0.1744 \\ 0.1744 & 0.2012 \end{bmatrix} \right).$$

Third, the probability of consumption reported on the 24HR was generated from a logistic model with age, level of education, and random intake component u_{1i} as the covariates. The intercept term was adjusted to -1.8 to obtain mean consumption probability of about 0.26 close to the estimate from NHANES.

Fourth, an indicator for a consumption status was randomly generated for each individual by comparing the generated consumption probability (π_{ij}) with a uniformly distributed random variable defined as $U_{ij} \in [0, 1]$. If $U_{ij} < \pi_{ij}$, a consumption is assumed on the 24HR, else a non consumption is assumed. For a non consumer, a zero intake value was assigned on the 24HR to denote non consumption. For consumers, however, the consumed amount was generated as follows. In the first simulation model, the consumed amount was generated following a generalized gamma distribution, with age, level of education, and random intake component u_{2i} as the covariates. To introduce heteroscedasticity, the parameter σ was generated following an exponential distribution with age as the covariate and using expression (7). In the second simulation model, the consumed amount was generated following a normal distribution on a Box–Cox transformed scale. The scale λ_R of the Box–Cox transformation was taken as 0.3394 and the within-person random component as $\varepsilon_{ij} \sim N(0, 0.5^2)$ as described in Section 4.3. Similarly, age, level of education, and random intake component u_{2i} were included as the covariates. The other simulation parameters were obtained from NHANES data.

Fifth, for each individual, 10 replicates of 24HR intake were generated and the true usual intake defined as the mean of the 10 replicate intake measurements for each individual.

Sixth, we generated normally distributed mercury level on a logarithmic scale as

$\log \text{merc}_i \sim N(1.49 + 1.34T'_i, \sigma = 1.15)$, where T'_i is on logarithmic scale to improve linearity. The model parameters were chosen to represent similar magnitude of association observed in NHANES.

Seventh, we applied the proposed and the standard calibration models to adjust for measurement error in intake, and the mean-24HR method that ignores measurement error in estimating the association between fish intake with the mercury level. In fitting each calibration model, we used the first two intake measurements reported in the 24HR as response values.

Lastly, we assessed the performance of the two calibration methods for various sizes of the simulation study, mainly by focusing on small to moderate study sizes. The performance of the two calibration models were further assessed by excluding the FFQ covariate from each part of the model. We ran 1000 simulations and assessed performance of the two calibration methods and the naive methods based on

1. Mean parameter estimate: $\bar{\hat{\beta}}_T = 1/N \sum_{k=1}^N \hat{\beta}_{T,k}$.
2. Relative bias in $\hat{\beta}_T$: $\text{Rel. bias}(\hat{\beta}_T) = \frac{1}{N} \sum_{k=1}^N (\hat{\beta}_{T,k} - \beta_T) / \beta_T$.
3. Empirical standard error of $\hat{\beta}_T$: $\text{Emp. SE}(\hat{\beta}_T) = \frac{1}{(N-1)} \sum_{k=1}^N (\hat{\beta}_{T,k} - \bar{\hat{\beta}}_T)^2$, defined as the standard deviation of $\hat{\beta}_T$ in N simulations and it quantifies the uncertainty in estimating β_T .
4. Mean squared error: $\text{MSE}(\hat{\beta}_T) = (\bar{\hat{\beta}}_T - \beta_T)^2 + (\text{Emp. SE}(\hat{\beta}_T))^2$, where β_T is the true association parameter, and $\hat{\beta}_{T,k}$ is the parameter estimate from the k -th simulation (Burton et al., 2006).

7 NHANES results

Figure 2 shows the distribution of usual fish intake estimated with the four methods: proposed calibration method, standard calibration method, mean-24HR method, and the FFQ method using NHANES 2003–2004 dataset. Regardless of the method used, a right-skewed distribution is shown for the estimated usual intake of fish. The distribution of usual intake estimated with the two calibration methods: the standard (dotted curve) and the proposed (solid curve) calibration methods are similar. Usual intake distribution estimated with the FFQ method (dashed curve), however, is shifted to the left of the distributions estimated with the two calibration methods. This suggests a potential systematic underestimation in the FFQ. Finally, usual intake distribution estimated with the 24HR (dashed dotted curve) differs substantially from the other estimated distributions due to high percentage of zero intake values in the mean-24HR method. This implies that fish is episodically consumed.

In comparing the goodness of fit of the two fitted calibration models, the proposed model fitted the calibration data better (AIC = 3826.6) than the standard model (AIC = 4375.8). Note, with the naïve methods, no calibration model is fitted, because the 24HR and FFQ-reported intakes are used to estimate the association with the outcome as if there were no measurement errors in the reported intakes.

Table 2 presents the mean (and standard deviation) of usual intake estimated with each of the four methods, the point estimate and the standard error of the regression coefficient, and R -squared from the linear outcome model that relates usual fish intake to blood mercury level in the NHANES 2003–2004 dataset. Among the four methods, the mean intake estimated with the FFQ method is the smallest. Further, the variability in usual intake estimates from the 24HR method is the greatest of all the estimates from the other methods. Again, the

mean intake estimated with the two calibration methods are of similar magnitude. Regression coefficient (that quantifies the association) adjusted for measurement error using either of the two regression calibration methods is about six times as large as the estimate obtained when usual intake is estimated by mean intake in the 24HR (the mean-24HR method), and about two times as large as the estimate obtained when usual intake is estimated by intake in the FFQ (the FFQ method). Specifically, with the mean-24HR method, the log-mercury level is estimated to increase by about $0.4 \mu\text{g/L}$ per ounce intake of fish, whereas with the FFQ method, the effect is estimated to increase by about $1.2 \mu\text{g/L}$. With both calibration methods, the effect is estimated to increase by about $2 \mu\text{g/L}$. The standard error estimates for the regression coefficient that is adjusted for measurement error using either of the two calibration methods are larger than the estimates obtained using the naïve methods. The observed large standard errors for the adjusted regression coefficient are due to the uncertainty in the calibration process. In regression calibration, usual intake is predicted using a regression function. The uncertainty in predicting usual intake results in larger standard error in the estimate of the adjusted regression coefficient. The naïve methods, however, do not suffer from this extra uncertainty as shown by smaller standard errors for the unadjusted coefficient estimates in Table 2. Based on the *R*-squared values for the fitted outcome models, usual intake that is calibrated with either of the two calibration methods fitted the outcome data better (larger *R*-squared value) than usual intake estimated with the naïve methods.

In Fig. 3, the regression coefficients estimated with the four methods are presented graphically. Again, the regression slopes estimated with the calibrated usual intake using the two calibration methods are similar and are steeper than the slopes estimated with the naïve methods. The slope estimate with the mean-24HR method is severely flattened.

8 Simulation results

Presented in Table 3 are the mean, relative bias, empirical standard error, and the mean squared error (MSE) for the regression coefficient estimate that quantifies the association between usual fish intake and mercury level in the blood in 1000 simulated datasets. True intake amount was generated following the generalized gamma distribution. The regression coefficient estimate that is adjusted for measurement error with either calibration methods is closer to the true coefficient than the unadjusted estimate. Based on the calibrated results, the coefficient estimate that is adjusted for measurement error with the proposed method is less biased (relative bias = 0.022) than the estimate adjusted with the standard method (relative bias = -0.069). Similar trend is observed in the precision of the coefficient estimates; for instance, the MSE of the coefficient estimate that is adjusted for measurement error with the proposed method is much smaller (MSE = 0.030) than that of the estimate that is adjusted for error with the standard method (MSE = 0.295). In contrast, the mean-24HR (naïve) method severely underestimates the true coefficient (relative bias = -0.934).

Table 4 presents the results when true usual intake amount was generated following a normal distribution on a Box-Cox transformed scale. Again, the regression coefficient estimates that are adjusted for measurement error with the two calibration methods are closer to the true value than the unadjusted estimate. Further, the proposed method seems to estimate the true

coefficient more precisely (smaller MSE and Emp. SE) than the standard method, but with slightly larger relative bias. Again, the naïve method severely underestimates the true coefficient.

In Table 5, the mean estimate, relative bias, empirical standard error, and MSE of regression coefficient estimates adjusted for measurement error with the two calibration methods for various sizes of simulated data are presented. True intake amount was generated following the generalized gamma distribution. Noteworthy, we encountered convergence issues with the standard method, mainly when study size was set as 250 or 500. As a result, extra simulations were run to obtain 1000 estimates of the regression coefficient where convergence was achieved in order to allow for proper comparisons with the results from the proposed method. For each method, the relative bias and MSE of the coefficient estimate decrease with increase in the study size. This means that the performance of the two calibration methods improved for larger study sizes and deteriorated for smaller study sizes.

We further assessed the performance of the two calibration methods with and without including FFQ covariate in each part of the model. From this comparison, the performance of each of the fitted calibration models deteriorated slightly when the FFQ covariate was excluded from each part of the model. For instance, relative bias and MSE increased by 4.5% and 5.6%, respectively, when FFQ covariate was excluded from each part of the proposed calibration model (Supporting Information Table S1).

9 Discussion and conclusion

In this paper, we proposed a flexible zero-augmented regression calibration to correct for measurement error in dietary usual intakes not consumed daily by everyone, before relating intake to a health outcome. The strength of the proposed calibration method is its ability to adjust for measurement error in intake, while handling the excess zeroes, skewness, and heteroscedasticity simultaneously without transforming the calibration response. The method further permits flexibility by modeling the correlation between the probability of consumption and the actual consumed amount. This correlation is common in nutritional research for dietary intakes not consumed daily such as fish and dark green vegetables. A further strength of the generalized gamma distribution assumed in part II of the proposed model is that it accommodates other simple standard distributions, such as gamma distribution shown in Section 3, and is able to approximate data that cannot be estimated with these simple distributions (Liu et al., 2010).

We applied the proposed calibration method to both real and simulated datasets and compared its performance with the standard calibration method and with the two methods that ignore measurement error in intake. In a simulation study, the proposed calibration method seemed to slightly outperform the standard method, when true intake amount was generated following either the generalized gamma distribution or a normal distribution on a Box–Cox transformed scale. The ability of the proposed method to perform well under these two distributional assumptions suggests its versatility.

From a nutritional point of view and for dietary intakes not consumed daily, we found that estimating diet–health outcome association with intake measurements from the 24HRs can be worse than with the intake measurements from the FFQ. Notably, in the 24HR, only food intake in the previous 24 hours is recorded, leading to substantial proportion of zeroes and large within-person variability for episodically consumed foods. In this study, 24HR was assumed as unbiased for true intake. In practice, however, this assumption might not hold, in which case objective biomarker measurements are preferred (Kipnis et al., 2003). Recovery biomarkers, however, do not exist for most dietary intakes such as fish intake, prompting the use of short-term self-report instruments such as 24HR. Nevertheless, in the absence of such a recovery biomarker, it has been shown that using 24HR can still provide a better approximation given substantial reporting bias in the FFQ as explained in Kipnis et al. (2009).

A potential limitation of the proposed calibration model concerns its complexity as there are many parameters to be estimated (see Section 4.2). The convergence of the model can be a serious issue, especially for smaller study sizes than those explored in this work, or for a response variable with very high proportion of zeroes. Fortunately, for the sample sizes explored in this study, we did not face convergence issues when using the proposed calibration model, both in real data analysis and in the simulation study. However, as noted before, we faced some convergence issues with the standard model in the simulation study, especially when sample size was set to either 250 or 500. A potential reason for the convergence problem with the standard method is that one has to approximate the inverse of the Box–Cox transformation in addition to other model parameters, including the random effects. Thus, for smaller studies, estimation methods, such as maximum likelihood might become intractable, posing a threat to the convergence of the proposed model. A potential remedy to the convergence problem could be to use more flexible Bayesian sampling-based estimation methods such as Markov chain Monte Carlo (Cooper et al., 2007; Rodrigues-Motta et al., 2015) or a more computationally efficient nonsampling-based Bayesian integrated nested Laplace approximation (Rue et al., 2009). Even though we assessed the goodness of fit of the proposed model using the AIC, an in-depth diagnostics is still needed with respect to the robustness and/or sensitivity of the proposed model to outliers, the degree of skewness, and heteroscedasticity of the response distribution. Though not the focus of this work, these aspects form a basis for further research.

Besides using the proposed method to adjust for measurement error, the proposed model and its subclasses have been used in a wide range of applications, such as in health economics (Liu et al., 2010), occupational epidemiology (Yau et al., 2002), modeling time-to-event outcomes (Chen et al., 2013), and in pharmacokinetics (Salway and Wakefield, 2008).

In conclusion, the proposed calibration method can correct for measurement error in episodically consumed foods, at least as good as the standard calibration method.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Chaozhi Zheng, Dr. Hilko van der Voet, and Prof. Hendriek Boshuizen of Biometris, Wageningen University, Netherlands for their useful and critical comments on the contents of the manuscript. We also thank the two anonymous reviewers and associate editor for their insightful comments that helped improve this work. This work was supported, in part, by the Yale Claude D. Pepper Older Americans Independence Center (P30AG021342). We acknowledge the use of NHANES 2003–2004 data.

References

- Agogo GO, van der Voet H, van't Veer P, Ferrari P, Leenders M, Muller DC, Sanchez-Cantalejo E, Bamia C, Braaten T, Knuppel S, Johansson I, van Eeuwijk FA, Boshuizen H. Use of two-part regression calibration model to correct for measurement error in episodically consumed foods in a single-replicate study design: EPIC case study. *PLoS One*. 2014; 9:e113160. [PubMed: 25402487]
- Agogo GO, van der Voet H, van't Veer P, van Eeuwijk FA, Boshuizen HC. Evaluation of a two-part regression calibration to adjust for dietary exposure measurement error in the Cox proportional hazards model: a simulation study. *Biometrical Journal*. 2016; 58:766–782. [PubMed: 27003183]
- Akaike H. Likelihood of a model and information criteria. *Journal of Econometrics*. 1981; 16:3–14.
- Burton A, Altman D, Royston P, Holder R. The design of simulation studies in medical statistics. *Statistics in Medicine*. 2006; 25:4279–4292. [PubMed: 16947139]
- Carroll RJ, Midthune D, Subar AF, Shumakovich M, Freedman LS, Thompson FE, Kipnis V. Taking advantage of the strengths of 2 different dietary assessment instruments to improve intake estimates for nutritional epidemiology. *American Journal of Epidemiology*. 2012; 175:340–347. [PubMed: 22273536]
- Carroll, RJ., Ruppert, D., Stefanski, LA., Crainiceanu, CM. *Measurement Error in Nonlinear Models*. Chapman & Hall/CRC; New York, NY: 2006.
- CDC. National Health and Nutrition Examination Survey. 2015. Available at <http://www.cdc.gov/nchs/nhanes/>
- Chen PC, Zhang JJ, Zhang RQ. Estimation of the accelerated failure time frailty model under generalized gamma frailty. *Computational Statistics & Data Analysis*. 2013; 62:171–180.
- Cooper NJ, Lambert PC, Abrams KR, Sutton AJ. Predicting costs over time using Bayesian Markov chain Monte Carlo methods: an application to early inflammatory polyarthritis. *Health Economics*. 2007; 16:37–56. [PubMed: 16981192]
- Efron, B., Tibshirani, R. *An Introduction to the Bootstrap*. Chapman & Hall; New York, NY: 1993.
- Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. *Journal of the National Cancer Institute*. 2011; 103:1086–1092. [PubMed: 21653922]
- Kaaks R, Ferrari P, Ciampi A, Plummer M, Riboli E. Uses and limitations of statistical accounting for random error correlations, in the validation of dietary questionnaire assessments. *Public Health Nutrition*. 2002; 5:969–676. [PubMed: 12638598]
- Kaaks R, Riboli E. Validation and calibration of dietary intake measurements in the EPIC project: methodological considerations. *International Journal of Epidemiology*. 1997; 26:S15–S25. [PubMed: 9126530]
- Keogh RH, White IR. A toolkit for measurement error correction, with a focus on nutritional epidemiology. *Statistics in Medicine*. 2014; 33:2137–2155. [PubMed: 24497385]
- Kipnis V, Carroll RJ, Freedman LS, Li L. Implications of a new dietary measurement error model for estimation of relative risk: application to four calibration studies. *American Journal of Epidemiology*. 1999; 150:642–651. [PubMed: 10490004]
- Kipnis V, Midthune D, Buckman DW, Dodd KW, Guenther PM, Krebs-Smith SM, Subar AF, Tooze JA, Carroll RJ, Freedman LS. Modeling data with excess zeros and measurement error: application to evaluating relationships between episodically consumed foods and health outcomes. *Biometrics*. 2009; 65:1003–1010. [PubMed: 19302405]
- Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, Bingham S, Schoeller DA, Schatzkin A, Carroll RJ. Structure of dietary measurement error: results of the OPEN

- biomarker study. *American Journal of Epidemiology*. 2003; 158:14–21. discussion 22–16. [PubMed: 12835281]
- Liu L, Strawderman RL, Cowen ME, Shih YCT. A flexible two-part random effects model for correlated medical costs. *Journal of Health Economics*. 2010; 29:110–123. [PubMed: 20015560]
- Manning W, Mullahy J. Estimating log models: to transform or not to transform. *Journal of Health Economics*. 2001; 20:461–494. [PubMed: 11469231]
- Manning WG, Basu A, Mullahy J. Generalized modeling approaches to risk adjustment of skewed outcomes data. *Journal of Health Economics*. 2005; 24:465–488. [PubMed: 15811539]
- Olsen M, Schafer J. A two-part random-effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association*. 2001; 96:730–745.
- Rodrigues-Motta M, Galvis Soto DM, Lachos VH, Vilca F, Baltar VT, Junior EV, Fisberg RM, Lobo Marchioni DM. A mixed-effect model for positive responses augmented by zeros. *Statistics in Medicine*. 2015; 34:1761–1778. [PubMed: 25682753]
- Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society Series B—Statistical Methodology*. 2009; 71:319–392.
- Salway R, Wakefield J. Gamma generalized linear models for pharmacokinetic data. *Biometrics*. 2008; 64:620–626. [PubMed: 17888039]
- Slimani N, Ferrari P, Ocke M, Welch A, Boeing H, van Liere M, Pala V, Amiano P, Lagiou A, Mattisson I, Stripp C, Engeset D, Charrondiere R, Buzzard M, van Staveren W, Riboli E. Standardization of the 24-hour diet recall calibration method used in the European Prospective Investigation into Cancer and Nutrition (EPIC): general concepts and preliminary results. *European Journal of Clinical Nutrition*. 2000; 54:900–917. [PubMed: 11114689]
- Slimani N, Kaaks R, Ferrari P, Casagrande C, Clavel-Chapelon F, Lotze G, Kroke A, Trichopoulos D, Trichopoulou A, Lauria C, Bellegotti M, Ocke MC, Peeters PH, Engeset D, Lund E, Agudo A, Larranaga N, Mattisson I, Andren C, Johansson I, Davey G, Welch AA, Overvad K, Tjønneland A, Van Staveren WA, Saracci R, Riboli E. European Prospective Investigation into Cancer and Nutrition (EPIC) calibration study: rationale, design and population characteristics. *Public Health Nutrition*. 2002; 5:1125–1145. [PubMed: 12639223]
- Souverein OW, Dekkers AL, Geelen A, Haubrock J, de Vries JH, Ocke MC, Harttig U, Boeing H, van 't Veer P, Consortium E. Comparing four methods to estimate usual intake distributions. *European Journal of Clinical Nutrition*. 2011; 65:S92–S101. [PubMed: 21731012]
- Stacy EW. A generalization of the gamma distribution. *Annals of Mathematical Statistics*. 1962; 33:1187–1192.
- Stacy EW, Mihram GA. Parameter estimation for a generalized gamma distribution. *Technometrics*. 1965; 7:349–358.
- Toozee J, Midthune D, Dodd K, Freedman L, Krebs-Smith S, Subar A, Carroll R, Kipnis V. A new method for estimating the usual intake of episodically consumed foods with application to their distribution. *Journal of American Diet Association*. 2006; 106:1575–1587.
- Willet, W. *Nutritional Epidemiology*. Oxford University Press; New York, NY: 1998.
- Yau KKW, Lee AH, Ng ASK. A zero-augmented gamma mixed model for longitudinal data with many zeros. *Australian & New Zealand Journal of Statistics*. 2002; 44:177–183.

Figure 1. Distribution of fish intake (in ounce per day) as reported on the two 24HRs: first (24HR1) and second (24HR2) and on the FFQ.

Figure 2. Distribution of usual fish intakes estimated with mean-24HR method (dashed-dotted curve), FFQ method (dashed curve), standard (dotted curve), and proposed calibration models (solid curve), NHANES 2003–2004.

Figure 3.

The regression slope for the association between usual fish intake and log-mercury level estimated using the two naïve methods: mean-24HR method (dashed-dotted line) and FFQ method (long-dashed line) and the proposed (solid line) and standard (dotted line) calibration models, NHANES 2003–2004.

Table 1

Summary measures for fish intake (ounce per day) distribution as reported in two 24HRs (24HR1 and 24HR2) and FFQ, NHANES 2003–2004.

Intake distribution	24HR1	24HR2	FFQ
Percentage of zeroes	85.04	86.04	3.24
Mean	0.395	0.404	0.215
Standard deviation	1.297	1.349	0.242
Coefficient of variation	3.283	3.338	1.123
Skewness	4.640	4.596	2.413
10th percentile	0.000	0.000	0.030
50th percentile (median)	0.000	0.000	0.130
90th percentile	1.297	1.223	0.520
Observations	1604	1547	1605

Table 2

The mean usual fish intake estimate $\bar{\hat{T}}$ (standard deviation), estimate and standard error of regression coefficient for the association between usual fish intake and log-mercury level using the two naïve methods that ignore measurement error and the two calibration methods, NHANES 2003–2004.

Calibration methods	$\bar{\hat{T}}(\text{SD})$	$\hat{\beta}_r$	$\text{SE}(\hat{\beta}_r)$	$R\text{-squared}$
Mean-24HR (naïve)	0.399 (0.986)	0.361	0.024	0.1042
FFQ method	0.215 (0.242)	1.196	0.100	0.0562
Standard model	0.357 (0.196)	2.178	0.434	0.1695
Proposed model	0.302 (0.196)	2.134	0.413	0.1585

Table 3

Mean, relative bias, empirical standard error, and mean squared error for the regression coefficient relating usual fish intake and mercury level in 1000 simulated data sets, with true regression simulated as $\log\text{mercury}_i \sim N(1.49 + 1.34T'_i, \sigma = 1.15)$; true consumed amount is generated from a generalized gamma model.

Calibration method	$\bar{\hat{\beta}}_T$	Mean rel. bias ($\hat{\beta}_T$)	Mean emp. SE ($\hat{\beta}_T$)	Mean MSE ($\hat{\beta}_T$)
True regression	1.342	0.002	0.031	0.001
Proposed model	1.369	0.022	0.172	0.030
Standard model	1.247	-0.069	0.535	0.295
Mean-24HR (naïve)	0.089	-0.934	0.006	1.564

Table 4

Mean, relative bias, empirical standard error, and mean squared error for the regression coefficient estimate that relates usual fish intake with mercury level in 1000 simulated data sets, true regression was simulated as $\log\text{meric}_i \sim N(1.49 + 1.34T'_i, \sigma = 1.15)$; consumed amount is generated from a linear model on a Box–Cox transformed scale.

Calibration method	$\bar{\hat{\beta}}_T$	Mean rel. bias ($\hat{\beta}_T$)	Mean emp. SE ($\hat{\beta}_T$)	Mean MSE ($\hat{\beta}_T$)
True regression	1.341	0.000	0.037	0.001
Proposed model	1.204	−0.101	0.127	0.034
Standard model	1.283	−0.043	0.232	0.057
Mean-24HR (naive)	0.086	−0.936	0.006	1.572

Table 5

Mean, relative bias, empirical standard error, and mean squared error for the regression coefficient relating usual fish intake and mercury level in 1000 simulated data sets by varying study size, with true regression simulated as $\log\text{mercury} \sim N(1.49 + 1.34T_i', \sigma = 1.15)$; true consumed amount is generated from a generalized gamma model.

Study size	Calibration method	$\bar{\beta}_T$	Mean rel. bias ($\hat{\beta}_T$)	Mean emp. SE ($\hat{\beta}_T$)	Mean MSE ($\hat{\beta}_T$)
250	Proposed model	1.231	-0.081	0.374	0.151
	Standard model	1.020	-0.239	0.617	0.483
500	Proposed model	1.365	0.019	0.358	0.128
	Standard model	1.145	-0.146	0.584	0.379
1000	Proposed model	1.374	0.025	0.241	0.059
	Standard model	1.139	-0.150	0.544	0.336
2000	Proposed model	1.369	0.022	0.172	0.030
	Standard model	1.247	-0.069	0.535	0.295