

RESEARCH ARTICLE

Open Access

Bayesian adjustment for measurement error in continuous exposures in an individually matched case-control study

Gabriela Espino-Hernandez¹, Paul Gustafson^{1*} and Igor Burstyn²

Abstract

Background: In epidemiological studies explanatory variables are frequently subject to measurement error. The aim of this paper is to develop a Bayesian method to correct for measurement error in multiple continuous exposures in individually matched case-control studies. This is a topic that has not been widely investigated. The new method is illustrated using data from an individually matched case-control study of the association between thyroid hormone levels during pregnancy and exposure to perfluorinated acids. The objective of the motivating study was to examine the risk of maternal hypothyroxinemia due to exposure to three perfluorinated acids measured on a continuous scale. Results from the proposed method are compared with those obtained from a *naïve* analysis.

Methods: Using a Bayesian approach, the developed method considers a classical measurement error model for the exposures, as well as the conditional logistic regression likelihood as the disease model, together with a random-effect exposure model. Proper and diffuse prior distributions are assigned, and results from a quality control experiment are used to estimate the perfluorinated acids' measurement error variability. As a result, posterior distributions and 95% credible intervals of the odds ratios are computed. A sensitivity analysis of method's performance in this particular application with different measurement error variability was performed.

Results: The proposed Bayesian method to correct for measurement error is feasible and can be implemented using statistical software. For the study on perfluorinated acids, a comparison of the inferences which are corrected for measurement error to those which ignore it indicates that little adjustment is manifested for the level of measurement error actually exhibited in the exposures. Nevertheless, a sensitivity analysis shows that more substantial adjustments arise if larger measurement errors are assumed.

Conclusions: In individually matched case-control studies, the use of conditional logistic regression likelihood as a disease model in the presence of measurement error in multiple continuous exposures can be justified by having a random-effect exposure model. The proposed method can be successfully implemented in WinBUGS to correct individually matched case-control studies for several mismeasured continuous exposures under a classical measurement error model.

Background

Measurement error refers to the variation of the observed measurement from the true value, and consists of two components, random error and systematic error. The first component, the random error, is caused by any factors that randomly affect the measurement across a sample, and usually arises from inaccuracy in a

measuring laboratory instrument or random fluctuations in the environmental conditions. The second error component, the systematic error, is caused by any factors that systematically affect the measurement across a sample, and can be attributed to non-random problems in the system of measurement (e.g. wrong use or improper calibration of the measurement instrument).

In many scientific areas where statistical analysis is performed, the problem of dealing with explanatory variables subject to measurement error is present. In particular, in epidemiologic studies, the explanatory

* Correspondence: gustaf@stat.ubc.ca

¹Department of Statistics, University of British Columbia, Vancouver, BC, Canada

Full list of author information is available at the end of the article

variables (or 'exposures') that reflect exposure to suspected risk factors associated with a disease (the outcome variable) are commonly measured with error. These errors can be either *differential* or *non-differential*, according to whether they depend on the values of other variables in the study, for instance the outcome variable [1,2]. As has been discussed by many authors [3-6], measurement error reduces power for detecting relationship between exposures and disease, and ignoring this error may bias the assessment of the association between health outcome and exposure variables. In particular, ordinary logistic regression can lead to biased estimates of odds ratios (ORs) when the covariates are subject to measurement error [7]. Researchers have proposed non-Bayesian methods to correct for measurement error in exposures in individually matched case-control studies. For instance, Guolo et al. [8] used conditional likelihood methods to correct for measurement error in a single continuous exposure using simulated data. These authors compared the performance of the likelihood methods with two other corrections techniques (regression calibration [7] and simulation-extrapolation (SIMEX) [9]), observing that the likelihood approach outperforms the alternative methods when a single continuous exposure is measured with error. McShane et al. [5] proposed a conditional scores procedure to correct for measurement error in some components of one or more continuous covariates. In that study, the authors treated the true covariates as fixed unknown parameters, which were removed from the likelihood by conditioning on a sufficient statistic and estimated together with the unknown parameters. However, the conditional scores procedure experienced convergence problems in the presence of large relative risks or when large measurement errors were considered. Also, conditional scores procedures are typically not very generalizable when data structures are changed even slightly. In addition, Liu et al. [6], Prescott and Garthwaite [10], and Rice [11] proposed Bayesian adjustments for misclassification of a binary exposure variable. Nevertheless, to our knowledge, very little attention has been given to measurement error in multiple continuous exposures in matched case-control studies, except for McShane et al. [5] whose procedure may be challenging numerically, and which is quite dependent on the settings of the problem.

Thus, in this paper, we develop a Bayesian method to correct for measurement error in multiple continuous exposures in individually matched case-control studies that may be generalized to different settings, where information regarding the measurement error variability is available from additional experiments. The methodology is illustrated using data from a study of association of perfluorinated acids (PFAs) with disruption of thyroid homeostasis in pregnant women [12]. PFAs are global

contaminants of human blood and environment [13]. The objective of the motivating study was to examine the risk of maternal hypothyroxinemia due to exposure to three PFAs measured on a continuous scale. No human study has previously examined the influence of PFAs on the development of hypothyroxinemia, but there are reports on the relationship between PFAs and thyroid hormones and thyroid disease. In a sample from the US general population from 1999 to 2006, both men and women exposed to some PFAs had higher prevalence of physician-diagnosed *thyroid disease* [14]. However, a small study in a highly contaminated community failed to find similar association [15], and two other studies also did not report associations [16,17]. Dallaire et al. [18] reported a mixture of negative and positive associations of thyroid hormones with PFAs. The extent to which measurement error may contribute to apparent heterogeneity among these reports is unknown, but it certainly should be considered as an explanation.

We start this paper by describing the data in the motivating example in detail, followed by derivation of an estimate of the random error variability from *percent recovery* experiments, description of the proposed Bayesian model and justification of conditional logistic disease model for measurement error correction. Next, application of the method is illustrated along with a sensitivity analysis of the impact on the results if greater-than-estimated random error was present. The proposed Bayesian method is implemented in WinBUGS software and inferences are compared to those drawn from a *naive* analysis, which ignores measurement errors in the exposures.

Methods

Data

The developed Bayesian method is illustrated using individually matched case-control data from a study of Chan et al. [12]. The objective of the study was to examine the risk of maternal hypothyroxinemia due to exposure to three PFAs: perfluorooctanoic acid (PFOA), perfluorooctane sulfonate (PFOS) and perfluorohexane sulfonate (PFHxS). Chan et al. [12] extracted PFAs from maternal sera samples from 271 pregnant women, aged 18 or older, who elected to undergo a second trimester prenatal "triple screen", delivered at 22 weeks gestation or more to live singletons without evidence of malformations, and were referred by a physician who made a least eight recommendations for the "triple screen" over the study period. The exposure variables were reported on a continuous scale, and censored/non-detectable values (about 5.4% of the total number of records) were recorded as half the value of the limit of detection. Concentrations of the PFAs were transformed to log-molar units, and it was seen that after this transformation the

measured exposures approximately follow a normal distribution. A quality control experiment was performed in order to assess the amount of error incurred in the measurement of the exposures. In this experiment, percentages of recovery were calculated for each exposure and the results revealed the presence of a random error in the measurements. Details of this procedure and results are presented in Appendix I.

Chan et al. [12] classified the subjects into cases or controls, based on the analysis of their thyroid stimulating hormone (TSH) and free thyroxine (T4) concentrations. The hypothyroxinemia cases correspond to women exhibiting normal TSH concentrations with no evidence of hyperthyroidism (between 0.15 and 4.0 mU/L) and free T4 in the 10th percentile (less than 8.8 pmol/L). Meanwhile the controls correspond to women with normal TSH concentrations but having free T4 concentrations between the 50th and 90th percentiles (between 12.0 and 14.1 pmol/L). Each case was matched to between one and three controls on the basis of two matching factors: maternal age at blood draw (± 3 years) and referring physician (a total of 29 physicians). Further details on the construction of the data can be found in Chan et al. [12].

In summary, the matched case-control data used to illustrate the Bayesian method to correct for measurement error contain information from 96 cases and 175 individually matched controls. For the purpose of this paper, it is assumed there is no misclassification of control/case status. In addition, the data contain, for each subject, the corresponding exposure to PFOA, PFOS and PFHxS, which are reported on a continuous scale in log-molar units and are assumed to be subject only to random measurement error. Moreover, four potential confounders which are precisely measured are reported: maternal age (years), maternal weight (pounds), maternal race (Caucasian and non-Caucasian) and gestational age (days). All potential confounders except for maternal race were reported on a continuous scale. The maternal age variable is retained despite its use as a matching factor, in case the matching is too coarse to fully eliminate confounding.

Measurement model

Generally, in observational studies, the vector of imprecise surrogate exposures $\underline{W}$ is commonly recorded, instead of $\underline{X}$ itself. Therefore, in order to understand the relationship between the disease risk and the explanatory variables $\underline{X}$, having data on Y and $\underline{W}$, it is necessary to account for measurement error in the exposures. In this paper, the attention is concentrated on the problem of having only random error, by assuming zero systematic error. However, the present methodology can be adapted to introduce the effect of a systematic error.

Assume the vector of independent surrogates $\underline{W}$ arises from a classical additive measurement error model, which can be expressed as

$$\underline{W} = \underline{X} + \underline{U}, \quad (1)$$

where $\underline{U}$ refers to the measurement error component. This classical model assumes the true exposures are recorded with an additive, independent error. In addition, it can be assumed the measurement error is *non-differential*, and *unbiased*. The assumption of *non-differential* measurement error refers to the fact that the distribution of the surrogate exposures depends only on the actual exposure variables and not on the response variable or other variables in the model. As a result, the conditional distribution of $(\underline{W}|\underline{X}, Y)$ is identical to the conditional distribution of $(\underline{W}|\underline{X})$. The *unbiased* assumption $E(\underline{U}|\underline{X}) = \underline{0}$ implies $E(\underline{W}|\underline{X}) = \underline{X}$. Typically, the measurement error component is also assumed to be normally distributed with constant variance, i.e. $\underline{U} \sim N_p(\underline{0}, \Sigma)$, where Σ is a diagonal matrix with the main diagonal entries given by $\Sigma_{pp} = \sigma_p^2$, for $p = 1, \dots, P$.

Under the stated assumptions, $\underline{W}|\underline{X}$ follows a P -dimensional multivariate normal distribution with a mean vector given by the vector of true exposures and a covariance matrix Σ , which in this case is known. Thus, the density corresponding to the measurement model is given by $\underline{W}|\underline{X}, \Sigma \sim N_p(\underline{X}, \Sigma)$. Therefore, under the assumptions of an individually matched case-control data, the density of the measurement model is given by

$$\prod_{i=1}^N \prod_{j=1}^{n_i} f(\underline{w}_{ij}|\underline{x}_{ij}, \Sigma), \quad (2)$$

where $\underline{w}_{ij}$ and $\underline{x}_{ij}$ correspond to the vector of surrogate and true exposures variables, respectively, for the j -th member of the i -th matched set, and N refers to the number of matched sets.

For the particular case of data used in the study on PFAs, the surrogate variables are measured concentrations of PFOA, PFOS, and PFHxS, which correspond to the exposures to the compounds reported on a continuous scale in log-molar units. Consequently, an additive measurement error model for the exposures in log-molar units translates into a multiplicative error structure, in which the corresponding error term is proportional to the true exposure in molar scale. In many epidemiological studies, positive explanatory variables are subject to this sort of measurement error. Using available validation data from the quality control procedure performed by Chan et al.[12], the covariance matrix Σ of the measurement model can be estimated. In Appendix I we present a statistical argument for estimating Σ from this particular form of quality control

data. The argument is based on the multivariate version of the delta method [19] and uses the estimated standard deviation of the percentages of recovery for the concentrations of the three compounds in parts-per-billion to obtain information about the incurred error in the measurement of the exposures.

Disease model

In order to describe a relationship between the true exposures and the probability associated to the response variable, it is necessary to specify a disease model. Since the study analysed in this paper involves matched sets, the conditional logistic regression likelihood is adopted.

Consider a study having N matched sets, such that the j -th member ($j = 1, \dots, n_i$) of the i -th set ($i = 1, \dots, N$) has P associated continuous exposures $\underline{X}_{ij} = (X_{ij1}, \dots, X_{ijP})^T$. In addition, let $\underline{Y}_i = (Y_{i1}, \dots, Y_{ini})^T$ be a vector of response variables associated to the i -th matched set, such that $Y_{ij} = 1$ for the cases and $Y_{ij} = 0$ for the controls. Without loss of generality, the subjects can be labelled such that $Y_{i1} = 1$ and $Y_{ij} = 0$, for $j = 2, \dots, n_i$. Thus, the underlying objective is to model the retrospective probabilities for the case (i.e. $P(\underline{X}_{i1} | Y_{i1} = 1)$), and the controls (i.e. $P(\underline{X}_{ij} | Y_{ij} = 0)$, for $j = 2, \dots, n_i$), which can be accomplished by using the conditional logistic regression model.

The conditional likelihood is obtained by conditioning on the number of cases in each matched set, i.e. conditioning on $\sum_{j=1}^{n_i} Y_{ij}$. In the particular case of individual matching, the number of cases is one; therefore, the conditional likelihood for the i -th matched set ($i = 1, \dots, N$) is given by [20,21]

$$L_i^C(\underline{\beta}) = f(\underline{y}_i = (1, 0, \dots, 0) | \underline{x}_i, \sum_{j=1}^{n_i} y_{ij} = 1) \\ = \frac{\exp(\underline{\beta}^T \underline{X}_{i1})}{\sum_{j=1}^{n_i} \exp(\underline{\beta}^T \underline{X}_{ij})}$$

where $\underline{\beta} = (\beta_1, \dots, \beta_P)^T$ corresponds to the log ORs associated with unit changes in each of the exposures. Under the assumptions of a matched case-control study, the full conditional likelihood is the product of $L_i^C(\underline{\beta})$ over the N strata or matched sets, which is

$$L^C(\underline{\beta}) = \prod_{i=1}^N L_i^C(\underline{\beta}). \quad (3)$$

The parameter $\underline{\beta}$ is assumed to be constant across matched sets, and it is the target of statistical inference.

Bayesian model

Consider a retrospectively collected matched case-control data where each case is matched to one or more controls based on suspected confounders as matching factors. Let Y be the response variable, such that $Y = 1$ for cases and $Y = 0$ for controls, let $\underline{X} = (X_1, \dots, X_P)^T$ be the P -dimensional vector of the true, latent, continuous exposures which are subject to measurement error, and let $\underline{W} = (W_1, \dots, W_P)^T$ be the P -dimensional vector of surrogate exposures.

The aim of this subsection is to develop a Bayesian method to understand the association between the vector of continuous exposures $\underline{X}$ and the probability of the response variable Y , after correcting for random measurement error in the exposures.

Under the Bayesian paradigm, the posterior density of the unknown quantities is given by

$$f(\underline{x}, \theta | \underline{w}, y) \propto f(\underline{x}, \underline{w} | y, \theta) f(\theta) \quad (4)$$

where θ refers to the vector of unknown parameters. The first term of the right hand side of (4) refers to the joint posterior distribution of the true exposures $\underline{X}$ and the surrogate variables $\underline{W}$. As will be shown in Appendix II, this term contains the densities of the measurement model, disease model and exposure model. Meanwhile, the second term corresponds to the prior distribution of the unknown parameters.

Exposure model

The conditional logistic regression model has been successfully applied in matched retrospective case-control studies, and the use of this procedure has been statistically justified using Bayesian (see for example [11,22,23]) and non-Bayesian (see for instance [20,21]) approaches. This justification is based on the fact that the likelihood term describing the distribution of the total number of cases within-stratum given the exposures can be discarded. The reason for this is that it does not provide information about the parameter of interest, since the likelihood is only a function of the unknown parameter β . However, this justification is no longer directly applicable when adjusting for measurement error in exposures, since the omitted likelihood term might also contain information about these exposures [6], i.e., the likelihood is a function of β and the unobserved exposures. As a result, the use of a conditional likelihood approach in the presence of measurement error in multiple continuous exposures has not been widely adopted.

We justify the use of conditional logistic regression likelihood as a disease model when adjusting for measurement error in an individually matched case-control study via a random-effect exposure model; details are presented in Appendix II. A different approach that

does not involve a random-effect exposure model is provided by Gulo et al. [8].

In order to describe the random-effect exposure model, we assume that the vector of exposures for the j -th subject from the i -th matched set follows a multivariate normal distribution around the vector of exposure means of the corresponding matched set. Moreover, since the vector of exposure means of a matched set is unknown, we assume that this vector follows a multivariate normal distribution centered on the across-set exposures means. That results in

$$\underline{X}_{ij} = \underline{\mu}_i + \underline{\gamma}_{ij}$$

and

$$\underline{\mu}_i = \underline{\bar{\mu}} + \underline{\lambda}_i$$

where $\underline{\gamma}_{ij} \stackrel{iid}{\sim} N_P(\underline{0}, V_W)$, and $\underline{\lambda}_i \stackrel{iid}{\sim} N_P(\underline{0}, V_B)$, such that $\underline{\gamma}_{ij}$ and $\underline{\lambda}_i$ are mutually independent. Also notice that the within-stratum covariance matrix V_W is assumed to be constant across matched sets. As a result, $\underline{X}_{ij}|\underline{\mu}_i \stackrel{iid}{\sim} N_P(\underline{\mu}_i, V_W)$, and $\underline{\mu}_i \stackrel{iid}{\sim} N_P(\underline{\bar{\mu}}, V_B)$. Thus, the density corresponding to the random-effect exposure model is given by

$$f(\underline{x}|\underline{\mu}, \underline{\bar{\mu}}, V_W, V_B) = \prod_{i=1}^N \prod_{j=1}^{n_i} f(\underline{x}_{ij}|\underline{\mu}_i, V_W) \times \prod_{i=1}^N f(\underline{\mu}_i|\underline{\bar{\mu}}, V_B). \quad (5)$$

It has been assumed that the vector of true exposures follows a P -dimensional multivariate normal distribution. However, in observational studies, exposures often have a skew distribution [24]. Therefore, it is important to keep in mind that incorrect model specification may lead to biased estimates. To overcome potential misspecifications, for the univariate case some authors [24-26] have proposed the use of flexible distributions to increase the robustness to model specification. However, implementation of such methods can be quite challenging in the context of multivariate exposures.

Joint posterior density

For the particular case of this paper, the data considered consist of $N = 96$ matched sets, such that the i -th matched set has $j = 1, \dots, n_i$ subjects, with $n_i \in \{2, 3, 4\}$. Thus one subject is the case per set and the remaining $n_i - 1$ subjects are the controls. Let $\theta = (\underline{\beta}, \underline{\mu}, \underline{\bar{\mu}}, V_W, V_B)$ be the vector of unknown parameters. Therefore, by (4) and (AII.3), and using densities in (2), (3) and (5), it follows that the posterior density of the unknown quantities can be expressed as

$$f(\underline{x}, \theta | \underline{w}, \gamma) \propto \prod_{i=1}^{96} \left\{ \prod_{j=1}^{n_i} f(\underline{w}_{ij} | \underline{x}_{ij}, \Sigma) \times \frac{\exp(\underline{\beta}^T \underline{X}_{i1})}{\sum_{j=1}^{n_i} \exp(\underline{\beta}^T \underline{X}_{ij})} \times \prod_{j=1}^{n_i} f(\underline{x}_{ij} | \underline{\mu}_i, V_W) \times f(\underline{\mu}_i | \underline{\bar{\mu}}, V_B) \right\} \times f(\underline{\beta}, \underline{\bar{\mu}}, V_W, V_B).$$

It is commonly assumed that the unknown parameters are independent of each other *a priori*, so that $f(\underline{\beta}, \underline{\bar{\mu}}, V_W, V_B) = f(\underline{\beta}) \times f(\underline{\bar{\mu}}) \times f(V_W) \times f(V_B)$. In order to implement a Bayes-Markov chain Monte Carlo (MCMC) inference, it is necessary to assume prior distributions for the unknown parameters. Proper prior distributions are assumed for all the parameters. Moreover, the corresponding hyperparameters are chosen so that the parameters reflect prior ignorance:

$$\underline{\beta} \sim N_P(\underline{0}, 10000I_P),$$

$$\underline{\bar{\mu}} \sim N_P(\underline{\mu}, 10000I_P),$$

$$V_W^{-1} \sim W_P(I_P, P + 2),$$

$$V_B^{-1} \sim W_P(I_P, P + 2),$$

where $W_P(R, b)$ indicates a P -dimensional Wishart distribution with a positive definite inverse scale matrix R , and b degrees of freedom. And I_P is an identity matrix of size P . For the particular case of the matched case-control data from the epidemiological study on PFAs, $P = 3$ and $\underline{\mu}$ is estimated using the across-set sample mean of the corresponding observed exposures.

Adjustment for additional confounders

Considering the possibility that confounding is only partially addressed by matching, further potential confounders can be introduced in the disease model. In general, potential confounders should also be included in the exposure model; however, for simplicity these confounders are not considered in our random-effect exposure model, keeping it as presented in equation (5). For the case of the PFA's data, this simplification might be justified by the fact that the exposures and the confounders exhibit small correlations (less than 0.18), so we do not expect the potential confounders to be very helpful in reconstructing the true exposures. In addition, due to the assumption of *non-differential* measurement error, the measurement model also remains as presented in equation (2).

Consider the situation where the j -th member of the i -th set has associated K potential confounding variables $\underline{Z}_{ij} = (Z_{ij1}, \dots, Z_{ijk})^T$ which are precisely measured. Therefore, the full conditional likelihood corresponding to the disease model in (3) can be rewritten as

$$L^C(\underline{\beta}, \underline{\delta}) = \prod_{i=1}^N \frac{\exp(\underline{\beta}^T \underline{X}_{i1} + \underline{\delta}^T \underline{Z}_{i1})}{\sum_{j=1}^{n_i} \exp(\underline{\beta}^T \underline{X}_{ij} + \underline{\delta}^T \underline{Z}_{ij})},$$

where $\underline{\delta} = (\delta_1, \dots, \delta_K)^T$ is the vector of parameters associated with the confounding effect.

Thus, the posterior density of the unknown quantities can be rewritten as

$$f(\underline{x}, \theta | \underline{w}, \underline{y}, \underline{z}) \propto \prod_{i=1}^{96} \left\{ \prod_{j=1}^{n_i} f(\underline{w}_{ij} | \underline{x}_{ij}, \Sigma) \times \frac{\exp(\underline{\beta}^T \underline{X}_{i1} + \underline{\delta}^T \underline{Z}_{i1})}{\sum_{j=1}^{n_i} \exp(\underline{\beta}^T \underline{X}_{ij} + \underline{\delta}^T \underline{Z}_{ij})} \times \prod_{j=1}^{n_i} f(\underline{x}_{ij} | \underline{\mu}_i, V_W) \times f(\underline{\mu}_i | \underline{\mu}, V_B) \right\} \times f(\underline{\beta}, \underline{\delta}, \underline{\mu}, V_W, V_B),$$

where $\theta = (\underline{\beta}, \underline{\delta}, \underline{\mu}, V_W, V_B)$ is the new vector of unknown parameters, and a proper and diffuse prior distribution is assumed for the parameter $\underline{\delta}$, by having $\underline{\delta} \sim N_K(\underline{0}, 10000I_K)$, where $K = 4$ for the particular case of the motivating example.

Results and Discussion

In this section, the proposed Bayesian method to correct for measurement error is illustrated using data from the study of Chan et al. [12]. Inferences drawn from a *naïve* analysis and an analysis correcting for measurement error are presented. The *naïve* analysis ignores error in exposure measurements, by pretending the observed exposures (PFOA, PFOS, and PFHxS) are precisely measured. Meanwhile, in the analysis accounting for measurement error, the surrogate exposures are corrected for random measurement error. Two models are considered in each analysis: a simple model assuming the only confounding is via matching factors, and a model adjusted by four further potential confounders (maternal age, maternal weight, maternal race, and gestational age). In summary, the results from four Bayesian models are compared: a simple model under the *naïve* analysis (N-S), an adjusted model by confounders under a *naïve* analysis (N-A), a simple model under a measurement error analysis

(ME-S), and an adjusted model by confounders under a measurement error analysis (ME-A).

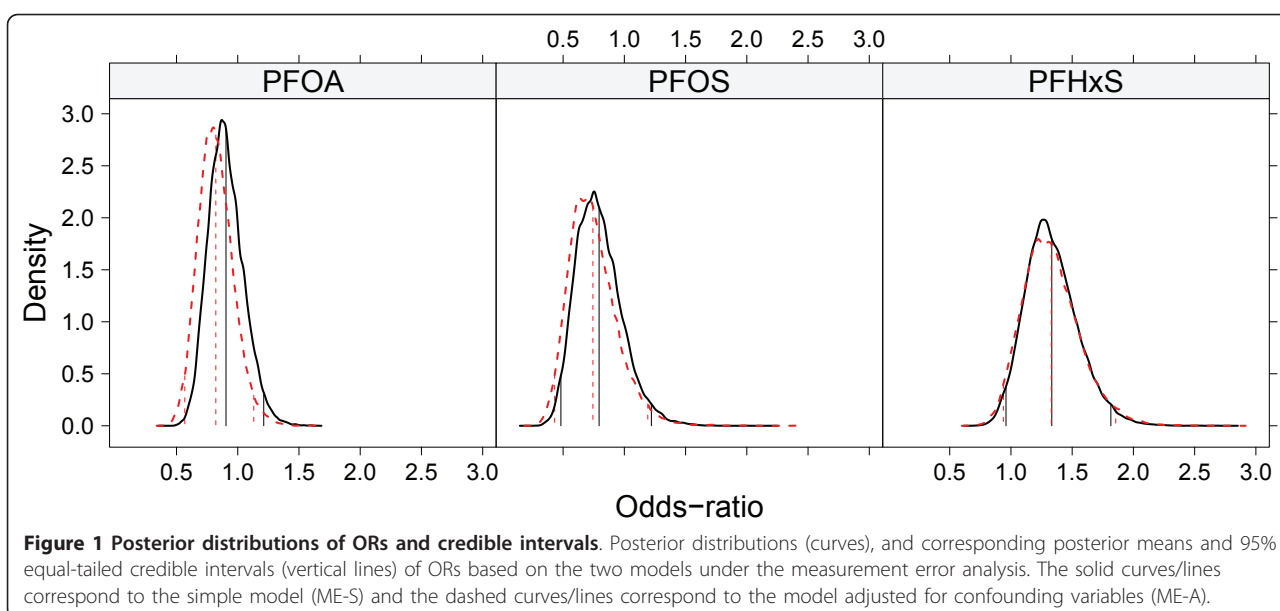
The models are implemented in WinBUGS software, version 1.4.3 [27], which is freely distributed and can be downloaded from the web [28]. Our WinBUGS code as available (Additional file 1). The analysis of the results was carried out using the statistical package R, version 2.11.1, which is also freely distributed on the web [29]. Two MCMC chains of length 55,000 were run for each model, using different initial values. The first 5,000 "burn-in" iterations were discarded from each chain and the last 50,000 MCMC iterations were used to perform Bayesian statistical inference. The computer running times on an Intel Core 2 Duo CPU at 2.10 GHz with 3.00 GB of RAM for N-S and N-A were approximately 1.5 and 4 minutes, respectively. Meanwhile, running times for ME-S and ME-A were about 9 and 13 minutes, respectively. The convergence to the posterior distributions and mixing of the two chains were assessed from the trace, autocorrelation, and the Gelman-Rubin convergence statistic plots. Moreover, under both types of analysis the estimated Monte Carlo standard errors of the posterior log ORs were smaller than 0.0026 for the simple models (N-S and ME-S) and smaller than 0.0030 for the models adjusted by the confounding variables (N-A and ME-A).

Posterior means and 95% equal-tailed credible intervals of the ORs obtained for the models under the *naïve* Bayesian analysis and the proposed Bayesian method to correct for measurement error are depicted in Table 1. This table indicates that under the two types of analysis, there is an absence of evidence for an association between the risk of maternal hypothyroxinemia and exposure to the PFAs. The point and credible intervals estimates of the ORs under both analyses are very similar, suggesting that a slight adjustment is manifested for the level of measurement error exhibited in the PFAs.

Table 1 Comparison of posterior means and credible intervals of the ORs

	OR	95% Cred. Int.	Adjusted OR	95% Cred. Int.
<i>Naïve analysis</i>				
PFOA	0.905	(0.661, 1.209)	0.828	(0.584, 1.127)
PFOS	0.802	(0.495, 1.214)	0.752	(0.445, 1.181)
PFHxS	1.315	(0.964, 1.755)	1.302	(0.934, 1.779)
<i>Measurement error analysis</i>				
PFOA	0.904	(0.656, 1.212)	0.821	(0.568, 1.131)
PFOS	0.794	(0.482, 1.221)	0.743	(0.431, 1.191)
PFHxS	1.333	(0.960, 1.816)	1.329	(0.938, 1.856)

Posterior means and 95% equal-tailed credible intervals of the ORs for the simple model and the model adjusted for confounding variables. Results are presented for the Bayesian *naïve* analysis and for the Bayesian method proposed to correct for measurement error.



However, *a priori*, there was not intuition that the adjustment would necessarily be slight.

Figure 1 presents the posterior densities and 95% credible intervals of the ORs for the simple model and the model adjusted by confounders, both under the measurement error analysis. Plots indicate there is no substantial association between the exposure to any PFA and maternal hypothyroxinemia. In addition, plots show a wider posterior distribution of the OR for the exposure to PFHxS, suggesting a bigger uncertainty in the risk associated to this exposure. Moreover, the posterior distributions for the simple model and the model after adjusting for the confounding variables are quite similar, in particular for the exposures PFOS and PFHxS, suggesting little or no further confounding effect. This conclusion is confirmed after calculating 95% equal-tailed credible intervals of the estimated parameter $\underline{\delta}$ (not shown).

A sensitivity analysis of the measurement error variability Σ estimated in Appendix I is carried out in order to validate the performance of the method. Figure 2 gives posterior means and 95% equal-tailed credible intervals for the ORs after increasing up to ten times the assumed measurement error variability for the exposures. Plots show that, for all the exposures, more substantial measurement error adjustments arise if larger measurement errors are assumed. Furthermore, the estimated ORs move in the anticipated directions, i.e., away from the null. However, in all cases the credible intervals widen, so that they still include the value of one, providing little evidence of any association between exposures and the outcome, regardless of the assumed measurement error magnitudes. The aforementioned

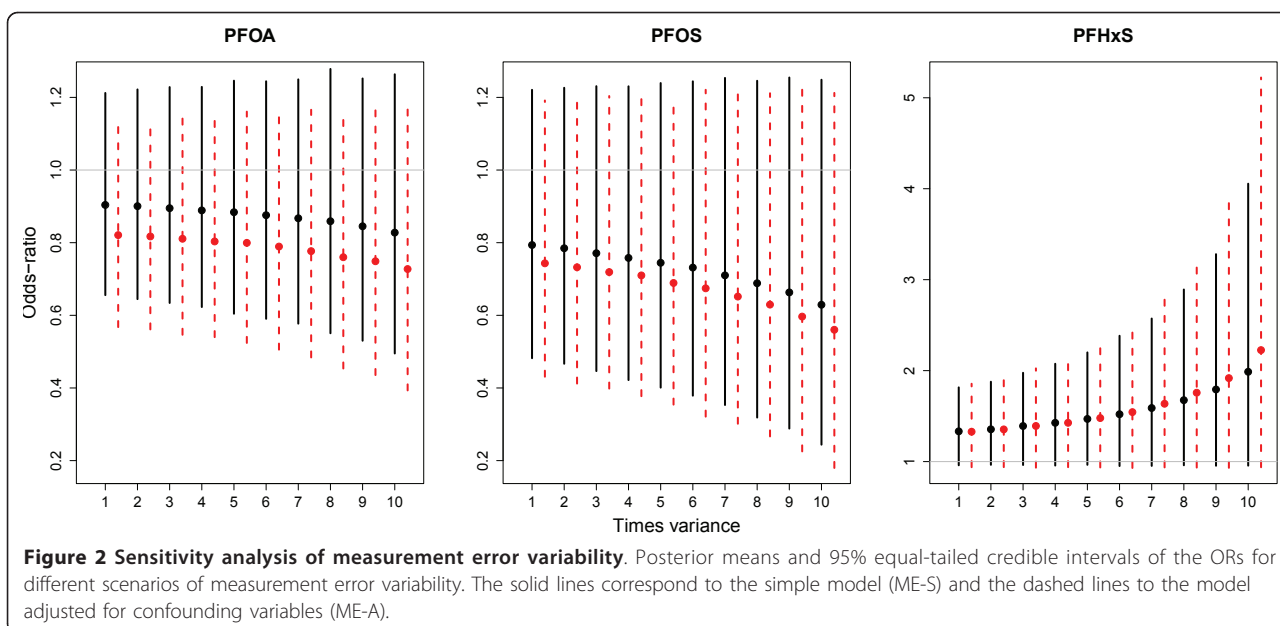
MCMC diagnostics indicated that MCMC convergence and mixing worsened slightly as the assumed measurement error magnitude increased. However, chains of length 55,000 were run with the first 5,000 interactions used as a “burn-in” period were still adequate.

Conclusions

We propose a Bayesian method to correct for measurement error in multiple continuous exposures for individually matched case-control studies. This method assumes a classical measurement model in order to account for random error in the exposures. It uses the conditional logistic regression likelihood as a disease model. We justify the use of this model in the presence of measurement error in the exposures by having a random-effect exposure model.

The proposed method can be implemented in WinBUGS software, which manages the computational complexity associated with likelihood-based approaches, to which Guolo et al. [8] referred. Moreover, as was pointed out by Guolo et al. [8], the likelihood-based methods, such as Bayesian and maximum-likelihood methods, perform well under different measurement error structures, can provide accurate inferential results, and outperform other corrections techniques (regression calibration and SIMEX). Furthermore, unlike the method proposed by McShane et al. [5] to correct for measurement error in continuous exposures, the Bayesian method proposed in this paper is neither prone to convergence errors nor highly dependent on the settings of a particular individually matched case-control study.

For the particular case of the study on PFAs, Bayesian inference of ORs indicates that little adjustment for



exposure measurement error is needed for the magnitude of error determined from the quality control experiment. However, bigger adjustments arise if larger measurement errors are assumed.

Some avenues for future research are suggested by our results. First, the method assumes a multivariate normal distribution on the exposures. However, it is important to keep in mind that a model misspecification may lead to biased estimates. In this context some authors have proposed the use of parametric and non-parametric flexible models. Nevertheless, some complications are involved in their implementations. For instance, Richardson et al. [24] proposed using a normal mixture model under a Bayesian approach and found that in the absence of validation data, their approach requires very strong priors on the mixture parameters to obtain reasonable estimates. Carroll et al. [25] suggested the use of a Bayesian approach in order to avoid the complicated implementation of the Estimation-Maximization (EM) algorithm under a traditional frequentist analysis once the normal mixture model is implemented into the likelihood. Furthermore, they advised to use partially proper priors in order to avoid improper posteriors. Guolo [26] suggested the use of skew-normal family of distributions as long as this distribution is a good approximation of the distribution of the unobserved exposures in the case-control sampling. Generally, however, the implementation of flexible exposure models for multivariate exposures remains challenging.

Second, we have not made explicit comparisons between our method and other methods. We have, however, considered implementation issues for our method

versus others. Particularly, we considered regression calibration techniques which impute best-guess exposure values and then plug these in to the disease model. While this is a simple procedure with some data formats, it would be no simpler than our method in the present format. The imputation involves estimating $E(X|W)$, which in turn requires estimating variance components from a multivariate random effect model applied to unbalanced data, in order to acknowledge variation between and within matched sets, in a similar fashion to [5]. Thus fitting a model similar to our exposure model is required, for which software options are somewhat limited. Moreover, regression calibration requires post-fitting adjustment of standard errors, say by bootstrapping, which would be very burdensome computationally in the present setting.

Finally, using available information from the quality control experiment performed on the PFA concentrations and the multivariate version of delta method, we present a statistical approach to estimate the measurement error variability. However, different assumptions and estimation methods can be developed in the presence of additional validation data or a different structure of quality control data. For instance, the complicated structure of the percent recovery experiments necessitated a 'plug-in' approach to dealing with the measurement error covariance matrix. Simpler data structures for informing the measurement error variance, such as a validation subsample, replicates, or an instrumental variable, would much more easily lend themselves to incorporating uncertainty about this covariance matrix as part of the overall Bayesian analysis.

Appendix I. Measurement error variability estimation

In reference to the epidemiological matched case-control study on PFAs, Chan et al. [12] performed a quality control procedure on the PFAs in ppb concentrations. Serum samples for subjects were divided into batches of approximately 16 per set for analysis of PFAs. Each set had a pooled sample consisting of: a paired sample of spiked serum (50 ppb of mixed standard in pooled serum) and unspiked serum (only pooled serum), besides a gold standard sample (50 ppb into methanol). Percentages of recovery were calculated by comparing the spiked concentration (i.e. difference between the paired spiked and unspiked samples) to the gold standard sample. The results showed that the standard deviations of the percent recoveries for PFAs in ppb concentrations were: 0.157 for PFOA, 0.139 for PFOS and 0.252 for PFHxS.

Let W_p^* , $W_{p,spiked}^*$ and $W_{p,gold}^*$ be the unspiked serum, spiked serum and gold standard corresponding to exposure p in ppb concentration, with $p \in \{1,2,3\} = \{PFOA, PFOS, PFHxS\}$. Using this notation, the percent recovery corresponding to exposure p is given by

$$Q_p = \frac{W_{p,spiked}^* - W_p^*}{W_{p,gold}^*}.$$

Let m_p be the factor used to convert molar units to ppb concentrations, corresponding to exposure p . Therefore, the percent recovery can be expressed as

$$Q_p = \frac{m_p * \exp(W_{p,spiked}) - m_p * \exp(W_p)}{m_p * \exp(W_{p,gold})}, \quad (AI.1)$$

where W_p , $W_{p,spiked}$ and $W_{p,gold}$ correspond to the unspiked serum, spiked serum and gold standard samples in log-molar concentrations, respectively.

Using the normality and homogeneity assumptions of the error component, the corresponding equation (1) for a particular exposure p is equivalent to

$$W_p = X_p + \sigma_p \varepsilon_p, \text{ with } \varepsilon_p \sim N(0, 1).$$

Therefore, it follows that

$$W_{p,spiked} = X_{p,spiked} + \sigma_p \varepsilon_{p,spiked}, \text{ with } \varepsilon_{p,spiked} \sim N(0, 1),$$

$$W_{p,gold} = X_{p,gold} + \sigma_p \varepsilon_{p,gold}, \text{ with } \varepsilon_{p,gold} \sim N(0, 1).$$

By substituting these three equations into (AI.1), it is possible to see that

$$Q_p = \frac{m_p * \exp(X_{p,spiked}) * \exp(\sigma_p \varepsilon_{p,spiked}) - m_p * \exp(X_p) * \exp(\sigma_p \varepsilon_p)}{m_p * \exp(X_{p,gold}) * \exp(\sigma_p \varepsilon_{p,gold})}. \quad (AI.2)$$

Moreover, according to the description of the samples in the quality control procedure $X_{p,spiked}$, $X_{p,gold}$ which

correspond to the true spiked serum and gold standard samples in log-molar concentrations, have the following underlying structures

$$\begin{aligned} m_p * \exp(X_{p,spiked}) &= m_p * \exp(X_p) + 50, \\ m_p * \exp(X_{p,gold}) &= 50. \end{aligned} \quad (AI.3)$$

Substitution of (AI.3) into equation (AI.2) yields

$$Q_p = \frac{(50 + c_p) * \exp(\sigma_p \varepsilon_{p,spiked}) - c_p * \exp(\sigma_p \varepsilon_p)}{50 * \exp(\sigma_p \varepsilon_{p,gold})},$$

where $c_p = m_p * \exp(X_p)$. Therefore, the percentage of recovery corresponding to exposure p is a twice-differentiable function of $T_p = (\varepsilon_p, \varepsilon_{p,spiked}, \varepsilon_{p,gold})^T$. Assuming $(\varepsilon_p, \varepsilon_{p,spiked}, \varepsilon_{p,gold})$ are independent, T_p follows a trivariate normal distribution with a mean vector of zeros and a covariance matrix equal to the identity matrix. Thus, based on the multivariate delta method, the variance of the percent recovery is given by

$$\begin{aligned} Var(Q_p) &\cong \nabla_{Q_p}(0)^T I_3 \nabla_{Q_p}(0) \\ &\cong \left[\left(\frac{c_p}{50} \right)^2 + \left(1 + \frac{c_p}{50} \right)^2 + 1 \right] * \sigma_p^2, \end{aligned} \quad (AI.4)$$

where is the gradient of $\nabla_{Q_p}(T_p)$. Using the gradient of Q_p results of the quality control procedure (standard deviations of the percentages of recovery for PFAs in ppb concentrations), and by taking c_p as the sample average of the ppb concentrations recorded for exposure across-sample, estimates for the measurement error variability for each exposure can be obtained as follows

$$\hat{\sigma}_p^2 \cong \frac{\hat{Var}(Q_p)}{\left[\left(\frac{\hat{c}_p}{50} \right)^2 + \left(1 + \frac{\hat{c}_p}{50} \right)^2 + 1 \right]}.$$

Using that information, the estimate of covariance matrix Σ in (2) is given by

$$\begin{aligned} \hat{\Sigma} &\cong \text{diag}(\hat{\sigma}_1^2, \hat{\sigma}_1^2, \hat{\sigma}_3^2) \\ &\cong \text{diag}(\hat{\sigma}_{PFOA}^2, \hat{\sigma}_{PFOS}^2, \hat{\sigma}_{PFHxS}^2) \\ &\cong \text{diag}(0.01180007, 0.007957311, 0.03042517). \end{aligned} \quad (AII.1)$$

Appendix II. Justification for conditional likelihood in matched case-control studies with measurement error in continuous exposures

Bayesian justifications for using conditional likelihood when actual exposure is observed are given by Rice [11,22,23], but the situation is less clear when the actual exposure is unobserved and treated as an unknown quantity inside the posterior distribution. Thus we

provide the following argument for using the conditional likelihood as a disease model, as long as the model for exposure acknowledges both across-stratum and within-stratum variation. For simplicity the argument is presented in the situation without confounders that vary within matched sets, i.e., all confounding is addressed via matching.

Under the Bayesian paradigm, for individually matched case-control data retrospectively collected and subject to measurement error, the joint posterior model of the true exposure and surrogate variables for a specific stratum (matched set) s is given by

$$f(\underline{y}, \underline{w} | \underline{y} = (1, 0, \dots, 0), s) = f(\underline{w} | \underline{y}, \underline{y} = (1, 0, \dots, 0), s) \times f(\underline{x} | \underline{y} = (1, 0, \dots, 0), s) \\ = f(\underline{w} | \underline{x}) \times f(\underline{x} | \underline{y} = (1, 0, \dots, 0), s).$$

The first term of the right hand side of (AII.1) is obtained under the assumption of *nondifferential* measurement error model. Regarding the second term, since $\sum_j Y_j$ depends explicitly on the values of the elements of the vector $\underline{Y}$, the distribution of $(\underline{X}, \underline{Y} = (1, 0, \dots, 0), S)$ is the same as the distribution of $(\underline{X}, \underline{Y} = (1, 0, \dots, 0), S, \sum_j Y_j = 1)$. Therefore, the retrospective probability in the second term of the right hand side of (AII.1) can be rewritten as

$$f(\underline{x} | \underline{y} = (1, 0, \dots, 0), s) = f\left(\underline{y} = (1, 0, \dots, 0) | \underline{x}, \sum_j y_j = 1, s\right) \times \\ \frac{f\left(\sum_j y_j = 1 | \underline{x}, s\right)}{f\left(\underline{y} = (1, 0, \dots, 0) | s\right)} \times f(\underline{x} | s). \quad (\text{AII.2})$$

Notice the distribution of $\underline{Y}$ is independent of the specific stratum, and it is mainly determined by the exposure variables. Therefore, by standard arguments, under a prospective logistic regression model with stratum-specific intercept, the probability of $\left(\underline{Y} = (1, 0, \dots, 0) | \underline{X}, \sum_j Y_j = 1, S\right)$ is simply the conditional likelihood term. On the other hand, arguably the dominant variation in the prospective distribution of $\left(\sum_j Y_j\right)$ given (X, S) will be with S rather than X , via the stratum-specific intercept. As a result, $f\left(\sum_j y_j = 1 | \underline{x}, s\right) \approx f\left(\sum_j y_j = 1 | s\right)$. Thus, (AII.2) can be approximated as

$$f(\underline{x} | \underline{y} = (1, 0, \dots, 0), s) \approx f\left(\underline{y} = (1, 0, \dots, 0) | \underline{x}, \sum_j y_j = 1\right) \times \\ \frac{f\left(\sum_j y_j = 1 | s\right)}{f\left(\underline{y} = (1, 0, \dots, 0) | s\right)} \times f(\underline{x} | s).$$

Since $Y_1, \dots, Y_{n_s} | S$ are *exchangeable*, it is possible to see that

$$\frac{f\left(\underline{y} = (1, 0, \dots, 0) | s\right)}{f\left(\sum_j y_j = 1 | s\right)} = f\left(\underline{y} = (1, 0, \dots, 0) | \sum_j y_j = 1, s\right) = 1/n_s.$$

Thus, the joint posterior density of the true exposure and surrogate variables for a specific stratum s can be expressed as

$$f(\underline{x}, \underline{w} | \underline{y} = (1, 0, \dots, 0), s) \propto f(\underline{w} | \underline{x}) \times f\left(\underline{y} = (1, 0, \dots, 0) | \underline{x}, \sum_j y_j = 1\right) \times f(\underline{x} | s), \quad (\text{AII.3})$$

where the conditional density of $(\underline{W} | \underline{X})$ corresponds to the measurement model. This describes how the surrogate vector of explanatory variables arises from the true values of $\underline{X}$. And the density $(\underline{X} | S)$ refers to the within-stratum density of the exposure model. The within-stratum density of the exposures can be implemented as a random-effect model. Therefore, the use of a conditional logistic regression model, when the exposures are measured with error, is justified by having a random-effect exposure model.

Additional material

Additional file 1: WinBUGS Code. Code used to perform the Bayesian adjustment for measurement error in a matched case-control study with multiple continuous covariates

Acknowledgements

The authors thank JW Martin, E Chan, F Bamforth and NM Cherry of the University of Alberta (Edmonton, Canada) for their contribution to generating data that motivated our work. This research was financially supported by the Canadian Institutes for Health Research (Funding Reference Number 62863).

Author details

¹Department of Statistics, University of British Columbia, Vancouver, BC, Canada. ²Department of Environmental and Occupational Health, School of Public Health, Drexel University, Philadelphia, PA, USA.

Authors' contributions

GEH developed and implemented the Bayesian method under supervision of PG. IB designed epidemiological study, oversaw its conduct and identified the need to better understand the impact of measurement error on *naïve* analysis. All authors contributed to writing of the manuscript, read its final form and approved it.

Competing interests

The authors declare that they have no competing interests.

Received: 5 January 2011 Accepted: 14 May 2011

Published: 14 May 2011

References

1. Thomas D, Stram D, Dwyer J: **Exposure Measurement Error: Influence on Exposure-Disease Relationships and Methods of Correction.** *Annual Review of Public Health* 1993, **14**:69-93.
2. Fosgate GT: **Non-differential measurement error does not always bias diagnostic likelihood ratios towards the null.** *Emerging Themes in Epidemiology* 2006, **3**:7.
3. Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM: *Measurement Error in Nonlinear Models. A Modern Perspective* Chapman & Hall; 2006.
4. Gustafson P: *Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments* Chapman and Hall/CRC Press; 2004.
5. McShane LM, Midthune DN, Dorgan JF, Freedman LS, Carroll RJ: **Covariate measurement error adjustment for matched case-control studies.** *Biometrics* 2001, **57**(1):62-73.
6. Liu J, Gustafson P, Cherry N, Burstyn I: **Bayesian analysis of a matched case-control study with expert prior information on both the misclassification of exposure and the exposure-disease association.** *Statistics in Medicine* 2009, **28**:3411-3423.
7. Rosner B, Spiegelman D, Willett WC: **Correction of Logistic Regression relative risk estimates and confidence intervals for measurement error: The case of multiple covariates measured with error.** *American Journal of Epidemiology* 1990, **132**(4):734-745.
8. Guolo A, Brazzale AR: **A simulation-based comparison of techniques to correct for measurement error in matched case-control studies.** *Statistics in Medicine* 2008, **27**:3755-3775.
9. Cook JR, Stefanski LA: **Simulation-Extrapolation Estimation in Parametric Measurement Error Models.** *Journal of the American Statistical Association* 1994, **89**(428):1314-1328.
10. Prescott GJ, Garthwaite PH: **Bayesian analysis of misclassified binary data from a matched case-control study with a validation sub-study.** *Statistics in Medicine* 2005, **24**(3):379-401.
11. Rice K: **Full-likelihood approaches to misclassification of a binary exposure in matched case-control studies.** *Statistics in Medicine* 2003, **22**(20):3177-3194.
12. Chan E, Burstyn I, Bamforth F, Cherry NM, Martin JW: **Perfluorinated acids and hypothyroxinemia in pregnant women.** *Environmental Research* 2011, **111**(4):559-564.
13. Hansen KJ, Clemen LA, Ellefson ME, Johnson HO: **Compound-Specific, Quantitative Characterization of Organic Fluorochemicals in Biological Matrices.** *Environmental Science & Technology* 2001, **35**(4):766-770.
14. Melzer D, Rice N, Depledge MH, Henley WE, Galloway TS: **Association between Serum Perfluorooctanoic Acid (PFOA) and Thyroid Disease in the U.S. National Health and Nutrition Examination Survey.** *Environ Health Perspect* 2010, **118**(5):686-692.
15. Emmett EA, Zhang H, Shofer FS, Freeman D, Rodway NV, Desai C, Shaw LM: **Community exposure to perfluorooctanoate: relationships between serum levels and certain health parameters.** *Journal of Occupational & Environmental Medicine* 2006, **48**(8):771-779.
16. Inoue K, Okada F, Ito R, Kato S, Sasaki S, Nakajima S, Uno A, Saijo Y, Sata F, Yoshimura Y, Kishi R, Nakazawa H: **Perfluorooctane Sulfonate (PFOS) and Related Perfluorinated Compounds in Human Maternal and Cord Blood Samples: Assessment of PFOS Exposure in a Susceptible Population during Pregnancy.** *Environ Health Perspect* 2004, **112**(11):1204-1207.
17. Peck JD, Robledo C, Neas B, Calafat AM, Sjodin A, Wild R, Cowan LD: **Effects of persistent organic pollutants on thyroid hormone levels during pregnancy [abstract].** *Annual Meeting of the Society for Epidemiologic Research June 24-27, 2008. Volume 167* Chicago, Illinois, USA: American Journal of Epidemiology; 2008, s103.
18. Dallaire R, Dewailly E, Pereg D, Dery S, Ayotte P: **Thyroid function and plasma concentrations of polyhalogenated compounds in Inuit adults.** *Environ Health Perspect* 2009, **117**(9):1380-1386.
19. Lachin JM: **A.3 Delta Method.** *Biostatistical Methods: The Assessment of Relative Risks* John Wiley and Sons; 2000, 455-457.
20. Breslow NE, Day NE: *Statistical methods in cancer research, Volume 1- The analysis of case-control studies* Lyon, France: International Agency for Research on Cancer (IARC); 1980.
21. Lachin JM: **7. Logistic Regression Models.** *Biostatistical Methods: The Assessment of Relative Risks* John Wiley and Sons; 2000, 247-316.
22. Rice K: **Equivalence between conditional and mixture approaches to the Rasch model and matched case-control studies, with applications.** *Journal of the American Statistical Association* 2004, **99**(466):510-522.
23. Rice K: **Equivalence between conditional and random-effects likelihoods for pair-matched case-control studies.** *Journal of the American Statistical Association* 2008, **103**(481):385-396.
24. Richardson S, Leblond L, Jaussent I, Green PJ: **Mixture Models in Measurement Error Problems, with Reference to Epidemiological Studies.** *Journal of the Royal Statistical Society Series A (Statistics in Society)* 2002, **165**(3):549-566.
25. Carroll RJ, Roeder K, Wasserman L: **Flexible Parametric Measurement Error Models.** *Biometrics* 1999, **55**(1):44-54.
26. Guolo A: **A Flexible Approach to Measurement Error Correction in Case-Control Studies.** *Biometrics* 2008, **64**(4):1207-1214.
27. Lunn DJ, Thomas A, Best N, Spiegelhalter D: **WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility.** *Statistics and Computing* 2000, **10**:325-337.
28. The BUGS Project. [http://www.mrc-bsu.cam.ac.uk/bugs/].
29. The Comprehensive R Archive Network (CRAN). [http://cran.r-project.org/].

Pre-publication history

The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2288/11/67/prepub

doi:10.1186/1471-2288-11-67

Cite this article as: Espino-Hernandez et al.: Bayesian adjustment for measurement error in continuous exposures in an individually matched case-control study. *BMC Medical Research Methodology* 2011 **11**:67.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

