

A Bayesian Multi-level model for estimating the diet/disease relationship in a multicenter study with exposures measured with error: the EPIC study

Pietro Ferrari^{1*}, Raymond J Carroll², Paul Gustafson³, Elio Riboli⁴

¹International Agency for Research on Cancer, Lyon, France

²Texas A&M University, College Station, USA

³Department of Statistics, University of British Columbia, Vancouver, Canada

⁴St Mary's Campus, Imperial College, London, UK

July 2, 2008

Abstract

In a multicenter study, the overall relationship between diet and cancer risk can be broken down into: [a] within-center relationships, which reflect the relationships at the individual level in each of the centers, and [b] a between-center relationship, which captures the association between exposure and disease risk at the aggregate level. In this work, we propose the use of a Bayesian multilevel model that takes into account the within- and between-center levels of evidence, using information at the individual and aggregate level. Correction for measurement error is performed in order to correct for systematic between-center measurement error in dietary exposure, and for attenuation biases in relative risk estimates within centers. The estimation of the parameters is carried out in a Bayesian framework using Gibbs sampling. The model entails a measurement, an exposure, and a disease component. Within the European Prospective Investigation into Cancer and Nutrition (EPIC) the association between lipid intake, assessed through dietary questionnaire and 24-hour dietary recall, and breast cancer incidence was evaluated. This analysis involved 21,534 women and 334 incident breast cancer cases from the EPIC calibration study. In this study, total energy intake was positively associated to breast cancer incidence at the aggregate level, while no effect was observed for fat. At the individual level, height was positively related to breast cancer incidence, while a weaker association was observed for fat. The use of multilevel models, which constitute a very powerful approach to estimating individual vs. aggregate levels of evidence should be considered in multicenter studies.

*Address for correspondence: International Agency for Research on Cancer, Lyon, France. *e-mail:* ferrari@iarc.fr

Introduction

In nutritional epidemiology, measurements of intakes of foods and/or nutrients contain a sizeable portion of random and systematic measurement errors [1]. It has been recognized that the relationship between diet and disease outcome produces biased estimates of associations, if measurement errors are not accounted for [2]. In addition, when the effects of different dietary factors are evaluated, measurement errors can bias the association of interest in any direction, depending on the particular error structure between measured exposures [3, 4], so that measurement error is likely to mask several examined relationships between diet and cancer incidence.

One solution to minimize the effect of measurement errors in dietary measurements is to increase the true overall heterogeneity of the dietary exposures studied, combining data from multiple studies conducted in populations with different dietary habits [5, 6]. This was the rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC), a multicenter cohort study on diet and cancer conducted in 10 Western European countries [7], where populations differing both in their dietary characteristics and disease rates are investigated to evaluate the relationship between diet and the risk of chronic diseases, particularly cancer [8]. In the EPIC study, the association between dietary exposures and disease risk can be assessed at (a) the individual level (within-center), by relating individual dietary intake measurements to disease outcome, and at (b) the group level (between-center), by comparing average dietary intakes with cohort specific incidence rates.

In a multicenter study, if within- and between-center estimates of exposure-disease relationships are similar, they corroborate each other and may be combined into a single overall summary estimate of relative risk [9, 10], as $\hat{\beta}_{pooled} = \hat{\beta}_B ICC + \hat{\beta}_W(1 - ICC)$, where the parameters express the ecological (B) and individual (W) association between exposure and disease, and ICC is an estimate of the intraclass correlation coefficient, which expresses the amount of between-center variability of exposure compared to the total.

In a multicenter study the aggregate and the individual level associations can be es-

estimated simultaneously [11]. A valuable solution for analyzing clustered data is to follow Bayesian integral approximation techniques, and the popular Markov chain Monte Carlo (MCMC) approach [12], which allows estimates of all model parameters of interest to be obtained. In this way the joint distribution of model parameters can be calculated as the limiting distribution of a specially constructed Markov chain. Samples from this distribution are obtained using simulation techniques such as Gibbs sampling [13].

Gibbs sampling is a popular approach for constructing an appropriate Markov chain by successively sampling from the joint conditional distributions. Theory shows that the full conditional distributions uniquely determine the joint distribution so that the limiting distribution of a Gibbs sampling approach is the joint distribution of interest [14].

In the present work the use of Bayesian estimation is described to evaluate the individual and the aggregate component of the association between lipid and energy intakes and the risk of breast cancer in the EPIC study.

Statistical Methods

Hierarchical modelling is a widely used approach to build complex models, by specifying a series of more simple conditional distributions. This is done by formulating the problems in terms of *conditional independence models*. In this way, models are generally characterized by the expression of a marginal model through a sequence of conditional models [15]. Using conditional independence assumptions, each variable in the model is related conditionally to only a few other variables. Complex problems are broken down into modular, possibly hierarchical components which have a relatively simple structure.

In extension to the seminal work of Clayton [16] and Richardson & Gilks [17] on measurement errors in a Bayesian framework, let $i = 1, \dots, n_k$ denote study subjects in center k , $k = 1, \dots, K$, and assume that

- Y_{ik} is a disease outcome indicator for individual i in center k ;

- X_{ik} is a vector of true unknown habitual dietary intakes;
- μ_{xk} is the a vector of true population means of habitual dietary intake in center k ;
- Z_{ik} is a vector of covariates measured without error.
- m_{zk} is the vector of population means of variables measured without error in center k .

While some risk factors, Z_{ik} , are measured without error, true dietary intake, X_{ik} , is unknown. The Bayesian model is specified through three structural components, which entail the formulation of a disease, a measurement error and an exposure model [17].

The Disease Model

In this study, the effect of energy (E) and lipid (L) intakes on breast cancer incidence was evaluated. In the context of community intervention trials, Gail et al. [18] suggest the possibility of allowing the odds ratio to depend upon the center, i.e.,

$$\begin{aligned}
& \text{pr}(Y_{ik} = 1 | X_{ik}, \mu_{xk}, Z_{ik}, m_{zk}) \\
&= H \left\{ \beta_1 + \mu_{xk}^T \beta_{D,B} + (X_{ik} - \mu_{xk})^T \beta_{D,W} + \gamma_{1k} + (X_{ik} - \mu_{xk})^T \gamma_{D,W,k} + \right. \\
&\quad \left. m_{zk}^T \beta_{C,B} + (Z_{ik} - m_{zk})^T \beta_{C,W} \right\}, \tag{1}
\end{aligned}$$

where $\mu_{xk} = (\mu_{xk,E}, \mu_{xk,L})$, and $X_{ik} = (X_{ik,E}, X_{ik,L})$, and H is the logistic distribution function. In model (1), the coefficients β 's are fixed effect parameters across centers, where the terms β_D s model the effect of dietary variables and β_C s model potential confounding factors. The terms $\beta_{.W}$ capture the within-center component of the association for individual level covariates, and $\beta_{.B}$ reflect the between-group component, for group level variables. Multilevel models allow the within- and between-center associations to be simultaneously evaluated [19, 20, 21]. The terms γ 's are random parameters which express deviation from

fixed effects and capture heterogeneity across centers, for which it is assumed that

$$\Omega_k = (\gamma_{1k}, \gamma_{D,W,k}^T)^T \sim \text{MVN}(0, \Sigma_\Omega).$$

The Measurement Error Model

Two types of dietary intake assessments were collected in EPIC: [i] for all study participants, a food frequency questionnaire (Q), or modified dietary history questionnaire for the assessment of individuals' habitual, long-term intake levels [22]; and [ii] a single highly standardized 24-hour recall (24-HDR) of actual food consumption during the previous day for a large representative sub-sample of the entire EPIC cohort [23].

In the case of two dietary exposures, the measurement error model for questionnaire measurement is defined as

$$\begin{aligned} Q_{ik} = & \alpha_1 + \mu_{xk}^T \alpha_{D,B} + (X_{ik} - \mu_{xk})^T \alpha_{D,W} + \lambda_{1k} + (X_{ik} - \mu_{xk})^T \lambda_{D,W,k} + \\ & m_{zk}^T \alpha_{C,B} + (Z_{ik} - m_{zk})^T \alpha_{C,W} + \epsilon_{ik} \end{aligned} \quad (2)$$

with $Q_{ik} = (Q_{ik,E}, Q_{ik,L})$. Dietary questionnaire measurements are assumed to be linearly related to true intake. The coefficients α_1 and $\alpha_{D,W}$ express constant and proportional scaling biases [6], whereas the residual terms model the random part of measurement errors and are assumed to be uncorrelated with true intake level X_{ik} [2]. In analogy with model (1), the λ 's are random parameters which model heterogeneity across centers, while the parameters α_C model the effect of confounders in the relation between questionnaire and true intake measurements. It is further assumed that

$$\begin{aligned} \Lambda_k &= (\lambda_{1k}, \lambda_{D,W,k}^T) \sim \text{N}(0, \Sigma_\Lambda) \\ \epsilon_{ik} &\sim \text{MVN}(0, \Sigma_\epsilon). \end{aligned}$$

The 24-HDR measurements, R_{ik} , measure actual food consumption during the previous day. These assessments are more work-intensive and costly, and they are assumed to provide

reference measurements of intake, being modelled as

$$\begin{aligned}
 R_{ik} &= X_{ik} + \nu_{ik} \\
 \nu_{ik} &\sim \text{MVN}(0, \Sigma_\nu).
 \end{aligned}
 \tag{3}$$

This model assumes that measurement errors in R_{ik} are totally random.

Correlated errors

The measurement error model in (2) and (3), for questionnaire and reference measurements respectively, assume that errors in the R and Q measurements are uncorrelated. In practice, this assumption may not be fully met, and some positive correlation between errors may occur, reflecting subjects tendencies to under- or over-report intake when different dietary assessment methods are used [4, 9, 10, 24]. In this work, two measurement error models have been considered, where:

- (a) errors in Q and R measurements, ϵ_{ik} and ν_{ik} , are assumed to be independent with covariances matrices Σ_ϵ and Σ_ν ;
- (b) ϵ_{ik} and ν_{ik} have a joint covariance matrix $\Sigma_{\nu\epsilon}$ (i.e. assuming conditional dependence between Q_{ik} and R_{ik} given $(X_{ik}, \mu_{xk}, Z_{ik}, \mu_{zk}) = (\mathbf{X}, \mathbf{Z})$). This condition is indicated in the text as dependence of $Q_{ik}|X_{ik}$ and $R_{ik}|X_{ik}$.

The Exposure Model

For true intake variability, in order to capture variation between-subjects within-center, as well as variation within-center, it assumed that

$$\begin{aligned}
 X_{ik} &\sim \text{MVN}(\mu_{xk}, \Sigma) \\
 \mu_{xk} &\sim \text{MVN}(0, \Sigma_x).
 \end{aligned}$$

Prior Distributions

The following prior distributions were assumed for the model parameters:

$$\begin{aligned}
 \Sigma^{-1} &\sim \text{Wishart}(D, N) \\
 \Sigma_x^{-1} &\sim \text{Wishart}(D_x, N_x) \\
 \Sigma_\epsilon^{-1} &\sim \text{Wishart}(D_\epsilon, N_\epsilon) \\
 \Sigma_\nu^{-1} &\sim \text{Wishart}(D_\nu, N_\nu) \\
 \mathcal{B}_0 &= (\beta_1, \beta_{D,W}^T, \beta_{D,B}^T, \beta_{C,W}^T, \beta_{C,B}^T)^T \text{ i.i.d. } \sim \text{N}(0, 20) \\
 \mathcal{A}_0 &= (\alpha_1, \alpha_{D,W}^T, \alpha_{C,W}^T)^T \text{ i.i.d. } \sim \text{N}(0, 20) \\
 \Sigma_\Omega^{-1} &\sim \text{Wishart}(D_\Omega, N_\Omega) \\
 \Sigma_\Lambda^{-1} &\sim \text{Wishart}(D_\Lambda, N_\Lambda).
 \end{aligned}$$

where D and N indicate a scale matrix and the rank of the Wishart distributions. Values of scale matrices used in this work are detailed in Appendix 1. To check the robustness of results, models with different scale matrices for the Wishart distributions were fit, particularly for Σ^{-1} , but virtually unchanged estimates were obtained. No prior information was available on each element of vectors \mathcal{A}_0 and \mathcal{B}_0 . Therefore each component was given vague and independent prior distributions.

Bayesian estimation of parameters aims at determining the posterior distribution of the parameters given the data. To obtain the marginal posterior distribution of model parameters the joint posterior distribution must be treated. In this work, Bayesian estimation was carried out with Gibbs sampling [13, 25], which generates samples from the joint posterior distribution of all parameters, given the data.

The joint distribution of all the variables involved in the problem can be written as the product of all the model conditionals and the prior distributions of model parameters, and it is detailed in Appendix 2.

Validation of results

The estimation of the error correlation terms leads to identifiability problems in a classical frequentist setting. In the absence of additional "objective" measurements, like, for example, recovery biomarkers [26], whose errors can be assumed to be independent of errors in self-reported measurements, standard approaches customarily assume that the errors between Q and R are uncorrelated.

In this work, a simulation study has been performed to evaluate the ability of the Bayesian models to estimate model parameters. For this purpose, using parameters estimated applying models (a) and (b) to actual data, two simulated EPIC data-sets were generated. The distributions of questionnaire and recall measurements were generated using relationships (2) and (3) respectively, and values in $\hat{\Sigma}_{\nu\epsilon}$ for the conditional dependence model, and $\hat{\Sigma}_{\nu}$ and $\hat{\Sigma}_{\epsilon}$ for the conditional independence model. Similarly, an outcome variable was generated according to a Bernoulli distribution and a logit relationship with the list of predictors in the disease model (1). Models (a) and (b) were applied to each of the two generated data-sets, and results compared. The aim was to assess the performance of the models to reproduce the findings under a fully-controlled scenario, where all model parameters are known.

The EPIC Data

In this work the relationship between total energy and lipid intakes and the incidence of breast cancer in the EPIC study was evaluated. A total of 21,534 subjects from the calibration sample, for whom both Q and R measurements were available, were included. In 6.3 years of average follow-up time, with a total of 135,205 person-years, the EPIC calibration sample developed 334 incident breast cancer cases. To avoid sparse numbers, study subjects were grouped in 17 centers located in 10 European countries: France (North-East, North-West, South, South-coast), Italy, Spain, United Kingdom, The Netherlands

(Bilthoven and Utrecht), Greece, Germany (Heidelberg and Potsdam), Sweden (Malmö and Umeå), Denmark (Aarhus and Copenhagen), and Norway.

Adjustment for confounding variables was performed at the individual and at the aggregate level in the disease model (1), using age, height and weight. In the measurement error model for Q_{ik} , inclusion of confounding factors was considered, but no appreciable effects were observed. In model (2) no other covariates than X_{ik} were included.

In model (1), $\beta_{D,W} = (\beta_2, \beta_4)$ and $\beta_{D,B} = (\beta_3, \beta_5)$ estimate the effect of energy and lipid intake, at the individual and aggregate level respectively. Similarly, the vector $\gamma_{D,W} = (\gamma_{2k}, \gamma_{4k})$ captures between-center heterogeneity of associations.

Dietary exposures were log-transformed. Before the analysis all variables were rescaled to have zero mean and variance close to one. This was done to optimize convergence of computations by reducing correlations between parameters while, at the same time, keeping interpretability of estimated quantities.

Once the model had converged, parameters estimates in the disease model were linearly rescaled to express a change in breast cancer risk for a 5% and 15% increase in total energy and fat intake respectively. In the EPIC population, a 15% increase in total fat (12 g/day = 108 Kcal/day) corresponds roughly to a 5% increase in total energy.

Results from the Bayesian model were compared to a more traditional approach, involving a two-step procedure. In extension of the calibration methodology introduced by Rosner [27], a random effects linear calibration model [28] was first used to compute predicted values of intake for all study participants. The regression calibration was carried out using the MIXED procedure of SAS software [29]. Parameter estimates were obtained by restricted maximum likelihood estimation. Then, a fixed effects logistic disease model, adjusted by the same list of confounding factors as in (1), was used to estimate the association between dietary exposures and breast cancer risk. The disease model was adjusted by centre, by means of indicator variables. This procedure allows only the individual component of the diet/disease relationship, after correction for measurement errors, to be computed.

The Bayesian model was carried out using the WINBUGS software [30]. Data were imported in WINBUGS, using a function that launches WINBUGS from R [31], developed by A. Gelman (available at <http://www.stat.columbia.edu/~gelman/bugsR>). For these analyses, three chains were run simultaneously. For each chain, initial values for parameters in models (1), (2) and (3) were randomly generated according to Normal distributions. Similarly, initial values for X_{ik} values were randomly generated according to a standardized multivariate normal distribution. A total of 45,000 iterations were obtained after a burn-in of 30,000.

Results

The center-specific sample sizes, number of breast cancer incident cases, the total and average Person-years (PY), and the average age are reported in Table 1. The French centers and Copenhagen provided the highest number of cancer cases, as a result of relatively long follow-up time (France), relatively aged participants (Copenhagen), and large sample size.

Standard Approach: Regression Calibration

In Table 2 observed and calibrated log-odds ratios are presented. The association between breast cancer incidence and dietary exposure assessed through Q measurements was negative for energy and positive for lipids. After calibration, the same relationships were stronger for both energy and lipid intakes. Height was positively and significantly associated to cancer incidence.

Validation study and Convergence Check

Results of the validation study for the conditional independence (a) and dependence (b) models are presented in Table 3, using simulated data assuming conditional independence. Parameter estimates based on posterior distributions of simulated data were overall com-

patible with parameter values based on actual data, particularly for variance estimates. It is noteworthy that $\hat{\Sigma}_{\nu\epsilon 12}$ and $\hat{\Sigma}_{\nu\epsilon 34}$ were close to zero, in line with the assumption of conditional independence, where no correlation between errors in Q and R measurements were assumed.

Model diagnostics showed that parameters attained convergence, although the chains showed large autocorrelation. After thinning (thin=5), less autocorrelation was observed but summary statistics of model parameters were virtually unchanged. The history plots suggest that, although individual chains did not mix well, the three chains overlap, thus indicating that the process is not too sensitive to starting values. The prior and posterior distributions of $\hat{\Sigma}_{11}$ showed an appreciable degree of updating during the analysis.

Bayesian conditional independence model

Means and associated 95% credible intervals of total ($\hat{\Sigma}$) and between-center ($\hat{\Sigma}_x$) variance-covariance matrices of true intakes estimated assuming conditional independence between $Q_{ik}|X_{ik}$ and $R_{ik}|X_{ik}$ are presented in Table 4. The estimated correlation coefficient between true lipids and energy intake is 0.845. The estimated within-centre variance components are 0.155 and 0.247, for energy and lipids respectively. After removing the variability due to measurement errors, model-based ICC values were higher than the observed quantities. Means and associated 95% credible intervals of the error variance-covariance matrix for R_{ik} ($\hat{\Sigma}_\nu$) and Q_{ik} ($\hat{\Sigma}_\epsilon$) measurements were shown in Table 5. High values of error correlations between energy and lipids were estimated, for both R (0.853) and Q (0.926) measurements.

Table 6 shows the means and associated 95% credible intervals for fixed and random effects in the EPIC measurement and disease models. To improve interpretability, square roots of variance components are reported.

The results of the disease model showed a strong relationship between energy intake and breast cancer incidence at the aggregate level (0.431, 95%CI 0.128,0.751), while a weaker negative association was observed at the individual level (-0.043, 95%CI -0.141,0.053). For

lipids, a modest negative association was observed at the aggregate level (-0.098, 95%CI -0.644,0.487), while a positive association was observed at the individual level (0.105, 95%CI: -0.123, 0.339). The associated variance components showed heterogeneity of associations across EPIC centers for energy (0.290 95%CI: 0.118,0.567), and lipid (0.316 95%CI: 0.142,0.574) intake on breast cancer incidence. Height was positively associated to breast cancer incidence (0.121, 95%CI: 0.023, 0.220), although a modest negative association was observed at the aggregate level.

Bayesian conditional dependence model

While the values of between-centre variability in Table 8 were very similar to values in table 5, the predicted within-centre variances were 0.018 for energy and 0.164 for lipids in Table 7. Individual level true correlation estimate between energy and lipids were equal to 0.604. The correlation between errors in Q_{ik} and R_{ik} (Table 8) were equal to 0.287 and 0.173, for energy and lipids respectively.

When assuming a non-null error correlation between questionnaire and 24-HDR measurements, results were rather similar to findings of conditional independence (table 9), with the exception of energy intake at the individual level, which was mildly positively associated to breast cancer risk (0.075 95%CI: -0.185,0.342).

Discussion

In multi-center cohort studies, relationships between exposure and disease risk can be studied at two, complementary levels: within- and between-center. In the context of the EPIC study, a Bayesian approach has been implemented to evaluate the relationship between diet and cancer using evidence at the individual and aggregate level, while correcting for measurement errors in dietary exposure.

The approach developed in this work consisted of three main components, (i) a disease,

(ii) a measurement, and (iii) an exposure model. This latter is particularly important, as it has been suggested that regression parameters in the disease model are rather sensitive to the specification of the parametric shape of the prior distribution of exposure [32]. Dealing simultaneously with the three components can be complex as it is necessary to treat the joint distribution of all quantities involved in the models, thus introducing serious statistical challenges [33]. Although it is very difficult to specify joint multivariate structures for complicated measuring processes, it may be much easier to factor such joint distributions into a series of conditional models, linked together in a hierarchical framework.

The model developed in this study had a natural application in the evaluation of the association between fat and energy and breast cancer incidence at the individual level was complemented by the evidence at the aggregate level, using information from 17 EPIC centers. Group-level energy was found to be positively associated with breast cancer incidence, while at the individual level the association was protective, although not statistically significant. Lipid intake was found to increase the risk of breast cancer at the individual level. However, these results were not able to reproduce the positive associations observed between fat and breast cancer mortality [34] at the aggregate level.

The inclusion of confounding factors might explain the discrepancy with a pure ecological analysis, together with the possibility that information from 17 centres only was too limited for inference at the aggregate level. Further research is needed within the EPIC study to evaluate whether intakes of specific types of fat are associated with the risk of breast cancer.

The possibility of estimating true variability of exposure in a Bayesian framework was evaluated. In a standard approach, regression calibration allows the parameters of the disease model to be corrected for measurement errors [27]. Regression calibration produces a shrunken distribution of predicted values, whose variability does not reflect the variance of true intake, but only the variance predicted by questionnaire measurements. In EPIC, where only one replicate of R measurements is available, no estimate of the true variation

of dietary exposure can be obtained using regression calibration [28].

The results of this study show that, once the variability attributed to random and systematic measurement error have been removed, the estimated within-centre variance of energy and lipids intake exposure has consistently diminished, for both conditional independence and dependence models.

The validation of study results provide robust evidence on the performance of the Bayesian model to accurately estimate parameters. However, the shrinking of true variability estimated in the exposure model is surprising, particularly for energy intake in the conditional dependence model. The within-centre component of total energy intake in the log scale were 0.015 and 0.002, for conditional independence and dependence model respectively. In the OPEN study, using doubly labelled water (DLW), the variance of true intake was estimated to be equal to 0.026 in women [35]. Within the Womens Health Initiative (WHI) study, energy intake was objectively determined using information from activity-related energy expenditure, resting metabolic rate and thermic effect of food in 102 postmenopausal women [36]. A value of the standard deviation of energy intake equal to 357 kcal was reported, while in our study the same values, after exponential transformation, were 224 and 74 kcal, for the two different models considered. Although part of these differences could be attributed to within-person systematic error not properly accounted in DLW and accelerometer measurements in the OPEN and WHI studies, the results of our work warn for caution, and further research is needed to evaluate the accuracy of the model, particularly the exposure component.

In this work, non-informative prior distributions were employed throughout, because the ability of the Bayesian model to provide accurate parameter estimates was evaluated. However, for future applications, evidence from informative studies could be incorporated into the model. In addition, evidence from the OPEN study suggests that 24-HDR measurements contain systematic component of measurement errors [35]. The evaluation of departures from the classic measurement error model for R_{ik} measurements needs to be

deeply investigated.

Although important issues of identifiability of parameters are raised, the Bayesian model showed overall good convergence. Following the formal statistical definition of identifiability, a model is said to be *identifiable* if different values of unknown quantities cannot correspond to the same distribution of *observable* data. Otherwise, the problem is said to be *non-identifiable*. Usually in identifiable situations, consistent parameter estimation can be achieved [37], in the sense that Bayesian, or maximum likelihood, estimates converge to true parameter values, as sample size increases. In non-identifiable situations, it may be that no amount of data can lead to the true parameter values, if the true distribution of observable data corresponds to more than one set of parameter values.

An interesting feature of Bayesian inference is that in some sense it 'works' whether or not one has parameter identifiability, and the mechanics of forming a posterior distribution and obtaining parameter estimates from the distribution can be carried out equally well in non-identifiable situations [37]. The performance of Bayesian inferences in nonidentifiable models has received little attention in the literature [38, 39, 40]. In some measurement error scenarios, Bayesian inference from a nonidentifiable model can be worthwhile, since the extent of non-identifiability is often modest when the available prior information is sufficiently good.

The perceived dilemma would be that assuming conditional independence of the two instruments given the true exposure is a strong assumption that is likely to be violated, but that relaxing this assumption would lead to a nonidentified, and therefore unusable model. In fact, however, the information content in a posterior distribution arising from a nonidentified model is highly problem-specific. In the present situation the actual-data and validation results both indicate that the data can be rather informative about the extent of correlation between R and Q given X. It is plausible to believe that some of this information can be derived from the disease component of the model. Hence it seems preferable to not assume the error correlation is zero a priori.

In many measurement error scenarios, additional knowledge or additional data are needed to gain identifiability. Some knowledge is needed about how measurement errors affect the observed quantities, i.e. how Q and R are distributed, given X . When dealing with dietary intakes, the evaluation of the measurement errors structure customarily makes use of different types of assessment methods that combine information from self-reported assessments (dietary questionnaires, 24-HDRs, diaries etc.), with more objective measurements (biochemical markers) to estimate true absolute intake [10, 35]. Due to logistic and financial limitations, the latter, however, are generally not available in large epidemiological investigations.

The EPIC study offers an ideal design to explore the use of advanced statistical techniques, because information on disease outcomes, as well as on different types of dietary measurements is available. Bayesian statistics offer the unique possibility to handle very difficult problems by providing tools to break down a complex structure into simpler components. This feature is doubly appealing because it allows all the components involved to become informative.

Acknowledgments

Carroll's research was supported by a grant from the National Cancer Institute (CA-57030).

Appendix 1

The following scale matrices D were used for the Wishart distributions: $D = \begin{bmatrix} 0.2 & 0.01 \\ 0.01 & 0.2 \end{bmatrix}$,

$$D_x = \begin{bmatrix} 0.8 & 0.14 \\ 0.14 & 0.8 \end{bmatrix}, D_\epsilon = \begin{bmatrix} 0.5 & 0.1 \\ 0.1 & 0.6 \end{bmatrix}, D_\nu = \begin{bmatrix} 0.6 & 0.1 \\ 0.1 & 0.8 \end{bmatrix}, D_{\nu\epsilon} = \begin{bmatrix} 0.6 & 0.06 & 0.04 & 0.03 \\ 0.06 & 0.5 & 0.02 & 0.05 \\ 0.04 & 0.02 & 0.4 & 0.03 \\ 0.03 & 0.05 & 0.03 & 0.5 \end{bmatrix},$$

$$D_{\Omega} = \begin{bmatrix} 0.2 & 0.06 & 0.02 \\ 0.06 & 0.1 & 0.01 \\ 0.02 & 0.01 & 0.15 \end{bmatrix}, \text{ and } D_{\Lambda} = \begin{bmatrix} 0.4 & 0.05 & 0.03 & 0.01 \\ 0.05 & 0.3 & 0.05 & 0.02 \\ 0.03 & 0.05 & 0.2 & 0.01 \\ 0.01 & 0.02 & 0.01 & 0.3 \end{bmatrix}.$$

Appendix 2

The following definitions are made for the disease model

$$\begin{aligned} G_{ik} &= (1, \mu_{xk}^T, (X_{ik} - \mu_{xk}^T), m_{xk}^T, (Z_{ik} - m_{xk}^T)^T)^T \\ &= \text{allows fixed effect covariates;} \\ L_{ik} &= (1, (X_{ik} - \mu_{xk}^T)^T)^T = \text{allows cross level covariates} \\ \text{pr}(Y_{ik} = 1 | Z_{ik}, X_{ik}, \text{center} = k) &= H(G_{ik}^T \mathcal{B}_0 + L_{ik}^T \Omega_k) \end{aligned}$$

Similarly, for the measurement error model, in a scenario without adjustment factors Z_{ik} ,

$$\begin{aligned} Q_{ik} &= \begin{pmatrix} Q_{ik,E} \\ Q_{ik,L} \end{pmatrix} \\ F_{ik}^T &= \begin{pmatrix} 1 & \mu_{xk,E} & (X_{ik,E} - \mu_{xk,E}) & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \mu_{xk,L} & (X_{ik,L} - \mu_{xk,L}) \end{pmatrix} \\ &= \text{covariates for fixed effects} \\ S_{ik}^T &= \begin{pmatrix} 1 & (X_{ik,E} - \mu_{xk,E}) & 0 & 0 \\ 0 & 0 & 1 & (X_{ik,L} - \mu_{xk,L}) \end{pmatrix} \\ &= \text{covariates for random effects} \\ Q_{ik} &= F_{ik}^T \mathcal{A}_0 + S_{ik}^T \Lambda_k + \epsilon_{ik} \\ \Delta_{ik} &= I(24\text{-HDR for person } i \text{ in center } k \text{ is measured}) \end{aligned}$$

The term Δ_{ik} is an indicator variable to express the presence of R_{ik} measurements (1=subject in the calibration sample, 0=otherwise).

The joint distribution of all the variables involved in the model where conditional independence between $Q_{ij}|X_{ij}$ and $R_{ij}|X_{ij}$ was assumed, reads, except for constants of proportionality, as:

$$\begin{aligned}
& \prod_{k=1}^K \prod_{i=1}^{n_k} ([H\{L_{ik}^T \Omega_k + G_{ik}^T \mathcal{B}_0\}]^{D_{ik}} \times [1 - H\{L_{ik}^T \Omega_k + G_{ik}^T \mathcal{B}_0\}]^{1-D_{ik}}) \\
& \times \prod_{k=1}^K |\Sigma_\epsilon|^{-n_k/2} \exp \left\{ -\frac{1}{2} \sum_{k=1}^{n_k} (Q_{ik} - F_{ik}^T \mathcal{A}_0 + S_{ik}^T \Lambda_k)^T \Sigma_\epsilon^{-1} (Q_{ik} - F_{ik}^T \mathcal{A}_0 + S_{ik}^T \Lambda_k) \right\} \\
& \times |\Sigma_\nu|^{-\sum_{k=1}^K \sum_{i=1}^{n_k} \Delta_{ik}/2} \exp \left\{ -\frac{1}{2} \sum_{k=1}^K \sum_{i=1}^{n_k} \Delta_{ik} (R_{ik} - X_{ik})^T \Sigma_\nu^{-1} (R_{ik} - X_{ik}) \right\} \\
& \times \prod_{k=1}^K \left[|\Sigma|^{-n_k/2} \exp \left\{ -\frac{1}{2} \sum_{i=1}^{n_k} (X_{ik} - \mu_{xk})^T \Sigma^{-1} (X_{ik} - \mu_{xk}) \right\} \right] \\
& \times |\Sigma_x|^{-K/2} \exp \left\{ -\frac{1}{2} \sum_{k=1}^K \mu_{xk}^T \Sigma_x^{-1} \mu_{xk} \right\} \\
& \times |\Sigma^{-1}|^{(N-p-1)/2} \exp \left\{ -\frac{1}{2} \text{trace}(D \Sigma^{-1}) \right\} \\
& \times |\Sigma_\epsilon^{-1}|^{(N_\epsilon - p_\epsilon - 1)/2} \exp \left\{ -\frac{1}{2} \text{trace}(D_\epsilon \Sigma_\epsilon^{-1}) \right\} \\
& \times |\Sigma_\nu^{-1}|^{(N_\nu - p_\nu - 1)/2} \exp \left\{ -\frac{1}{2} \text{trace}(D_\nu \Sigma_\nu^{-1}) \right\} \\
& \times |\Sigma_x^{-1}|^{(N_x - p_x - 1)/2} \exp \left\{ -\frac{1}{2} \text{trace}(D_x \Sigma_x^{-1}) \right\} \\
& \times |\Sigma_\Omega^{-1}|^{K/2} \exp \left\{ -\frac{\sum_{k=1}^K \Omega_k^T \Sigma_\Omega^{-1} \Omega_k}{2} \right\} \\
& \times |\Sigma_\Lambda^{-1}|^{K/2} \exp \left\{ -\frac{\sum_{k=1}^K \Lambda_k^T \Sigma_\Lambda^{-1} \Lambda_k}{2} \right\} \\
& \times |\Sigma_\Omega^{-1}|^{(N_\Omega - p_\Omega - 1)/2} \exp \left\{ -\frac{1}{2} \text{trace}(D_\Omega \Sigma_\Omega^{-1}) \right\} \\
& \times |\Sigma_\Lambda^{-1}|^{(N_\Lambda - p_\Lambda - 1)/2} \exp \left\{ -\frac{1}{2} \text{trace}(D_\Lambda \Sigma_\Lambda^{-1}) \right\} \\
& \times \exp \left(-\frac{\mathcal{B}_0^T \Sigma_{\mathcal{B},p}^{-1} \mathcal{B}_0}{2} - \frac{\mathcal{A}_0^T \Sigma_{\mathcal{A},p}^{-1} \mathcal{A}_0}{2} \right).
\end{aligned}$$

One single expression for $\Sigma_{\nu\epsilon}$ replaces terms for Σ_ϵ and Σ_ν in a model where conditional dependence between $Q_{ij}|X_{ij}$ and $R_{ij}|X_{ij}$ is assumed.

References

- [1] Willett WC. An overview of issues related to the correction of non-differential exposure measurement error in epidemiologic studies. *Stat Med* 1989; **8(9)**: 1031-40.
- [2] Carroll RJ, Ruppert D, Stefanski LA, Crainiceanu CM. Measurement error in Nonlinear Models: A Modern Perspective. London: Chapman and Hall, CRC Press, Second Edition, 2006.
- [3] Day NE, Ferrari P. Some methodological issues in nutritional Epidemiology. *IARC Sci Publ* 2002; **156**: 5-10.
- [4] Day NE, Wong MY, Bingham S et al. Correlated measurement error—implications for nutritional epidemiology. *Int J Epidemiol* 2004; **33**: 1373-81.
- [5] Wynder EL, Hebert JR. Homogeneity in nutritional exposure: an impediment in cancer epidemiology. *J Natl Cancer Inst* 1987; **79(3)**: 605-7.
- [6] Kaaks R, Riboli E, Estve J. Estimating the accuracy of dietary questionnaire assessments: validation in terms of structural equation models. *Stat Med* 1994; **13**: 127-42.
- [7] Riboli E. Nutrition and cancer: background and rationale of the European Prospective Investigation into Cancer and Nutrition (EPIC). *Ann Oncol* 1992; **3**: 783-91.
- [8] Riboli E, Kaaks R. The EPIC Project: rationale and study design. European Prospective Investigation into Cancer and Nutrition. *Int J Epidemiol* 1997; **26** Suppl 1: S6-14.
- [9] Kaaks R, Plummer M, Riboli E, et al. Adjustment for bias due to errors in exposure assessments in multicenter cohort studies on diet and cancer: a calibration approach. *Am J Clin Nutr* 1994; **59** Suppl 1: 245S-50S.
- [10] Plummer M, Clayton D, Kaaks R. Calibration in multi-centre cohort studies. *Int J Epidemiol* 1994; **23**: 419-26.

- [11] Navidi W, Thomas D, Stram D, Peters J. Design and analysis of multilevel analytic studies with applications to a study of air pollution. *Environ Health Perspect* 1994; **102** Suppl 8: 25-32.
- [12] Gilks WR, Richardson S, Spiegelhalter DJ. Markov Chain Monte Carlo in Practice. London: Chapman & Hall, 1996.
- [13] Geman S and Geman D. Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images IEEE Trans. *Pattern Analysis and Machine Intelligence*, pp 721-71, vol 6, 7, 1984.
- [14] Besag J. Spatial interaction and the statistical analysis of lattice systems. *J R Stat Soc Ser B*1974; **36**: 192-236.
- [15] Zaslavsky A. Hierarchical Bayesian Modeling in Subjective and Objective Bayesian Statistics: Principles, Models, and Applications, 2nd Edition. Press J, 2002.
- [16] Clayton DG. Models for the analysis of cohort and case-control studies with inaccurately measured exposures. In: Statistical models for longitudinal studies of health (Dwyer JH, Lippert P, Feinleib M, Hoffmeister H, eds). Oxford: Oxford University Press, 1992; 301-31.
- [17] Richardson S, Gilks WR. Conditional independence models for epidemiological studies with covariate measurement error. *Stat Med* 1993; **12(18)**: 1703-22.
- [18] Gail MH, Mark SD, Carroll RJ, Green SB, Pee D. On design considerations and randomization-based inference for community intervention trials. *Stat Med*. 1996; **15(11)**: 1069-92. Review.
- [19] Snijders TAB, Bosker R. Multilevel analysis: an introduction to basic and advanced multilevel modelling. London, United Kingdom: Sage Publications, 1999.

- [20] Sheppard L. Insights on bias and information in group-level studies. *Biostatistics*. 2003; **4**: 265-78.
- [21] Begg MD, Parides MK. Separation of individual-level and cluster-level covariate effects in regression analysis of correlated data. *Stat Med* 2003; **22**: 2591-602.
- [22] Riboli E, Hunt KJ, Slimani N, et al. European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* 2002; **5**: 1113-24.
- [23] Slimani N, Ferrari P, Ocke M, et al. 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. *Eur J Clin Nutr* 2000; **54**: 900-17.
- [24] 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. *Am J Epidemiol* 1999; **150**: 642-51.
- [25] Gelfand AE, Hills SE, Racine-Poon A, and Smith AFM. Illustration of Bayesian inference in normal data models using Gibbs sampling. *JASA* 1990; **85**: 972-85.
- [26] Kaaks R, Ferrari P, Ciampi A, et al. Uses and limitations of statistical accounting for random error correlations in the validation of dietary questionnaire assessments. *Public Health Nutr* 2002; **5**: 969-76.
- [27] Rosner B, Willett WC, Spiegelman D. Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error. *Stat Med* 1989; **8**: 1051-69 (Discussion 1071-3).
- [28] Ferrari P, Kaaks R, Fahey MT, et al. Within and between cohort variation in measured macro-nutrient intakes in the EPIC study taking account of measurement errors. *Am J Epidemiol* 2004; **160(8)**: 814-22.

- [29] SAS Institute. SAS Online Doc, Version 8. Cary, NC: SAS Institute Inc., 2002.
- [30] Spiegelhalter DJ, Thomas A, Best NG, Gilks WR. Bayesian inference under Gibbs sampling, Cambridge, UK: Med Res Counc Biostat Unit, 1995.
- [31] R: A language and environment for statistical computing, 2004.
- [32] Richardson S, Leblond L. Some comments on misspecification of priors in Bayesian modelling of measurement error problems. *Stat Med* 1997; **16**: 203-13.
- [33] Wikle CK. Hierarchical models in environmental science. *International Statistical Review* 2003; **71**: 181-99.
- [34] Carroll KK. Experimental evidence of dietary factors and hormone-dependent cancers. *Cancer Res* 1975; **35**: 3374-83.
- [35] Kipnis V, Subar AF, Midthune D, et al. Structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol* 2003; **158**: 14-21.
- [36] Horner NK, Patterson RE, Neuhaus ML, et al. Participant characteristics associated with errors in self-reported energy intake from the Women's Health Initiative food-frequency questionnaire. *Am J Clin Nutr* 2002; **76(4)**: 766-73.
- [37] Gustafson P. Measurement Error and Misclassification in Statistics and Epidemiology. Chapman & Hall. Vancouver, Canada, 2004.
- [38] Neath AA, Samaniego FJ. On the Efficacy of Bayesian Inference for Nonidentifiable Models. *The American Statistician* 1997; **51**: 325-32.
- [39] Gustafson P, Le ND, Saskin R. Case-control analysis with partial knowledge of exposure misclassification probabilities. *Biometrics* 2001; **54**: 598-609.

- [40] Gustafson P. On Model Expansion, Model Contraction, Identifiability and Prior Information: Two Illustrative Scenarios Involving Mismeasured Variables. *Statistical Science* 2005; **20**: 111-40.

Table 1: Characteristics of the calibration sample used in the present analysis.

Center	sample size	cancer ^a cases	PY ^b		Age ^c	
			Sum	Mean	Mean	Std
North-East of France	1824	48	15603	8.6	54.0	6.9
North-West of France	567	8	4851	8.6	53.8	6.6
South of France	1297	43	11021	8.5	54.2	6.8
South coast of France	560	13	4772	8.5	54.6	6.6
Italy	2354	29	14551	6.2	52.7	7.3
Spain	1370	14	9323	6.8	50.8	8.1
United Kingdom	698	5	3526	5.1	54.2	9.3
Bilthoven	1336	14	5315	4.0	43.7	11.2
Utrecht	1699	35	11589	6.8	58.5	6.2
Greece	1372	2	5022	3.7	54.9	11.1
Heidelberg	1009	9	4914	4.9	49.9	8.5
Potsdam	963	7	5816	6.0	52.8	8.5
Malmö	1381	25	10464	7.6	58.8	7.6
Umeå	1470	26	11898	8.1	51.4	8.6
Aarhus	492	10	2889	5.9	55.8	4.3
Copenhagen	1448	40	8436	5.8	57.0	4.3
Norway	1694	6	5205	3.1	48.2	4.3
Total	21534	334	135204	6.3	53.1	8.6

^aBreast cancer incident cases

^bPerson-years

^cAge at recruitment

Table 2: Parameter estimates and associated SE of the EPIC disease model, before and after calibration.

exposure		Original ^a		Calibrated ^b	
		estimate	SE	estimate	SE
intercept	$\hat{\beta}_1$	-5.631	0.414	-5.613	0.413
energy ^c	$\hat{\beta}_2$	-0.017	0.020	-0.055	0.045
mean energy	$\hat{\beta}_3$	—	—	—	—
lipids ^d	$\hat{\beta}_4$	0.043	0.049	0.133	0.100
mean lipids	$\hat{\beta}_5$	—	—	—	—
age ^e	$\hat{\beta}_6$	0.144	0.040	0.147	0.040
mean age	$\hat{\beta}_7$	—	—	—	—
weight ^f	$\hat{\beta}_8$	0.027	0.026	0.024	0.027
mean weight	$\hat{\beta}_9$	—	—	—	—
height ^g	$\hat{\beta}_{10}$	0.116	0.050	0.124	0.055
mean height	$\hat{\beta}_{11}$	—	—	—	—

^aAll parameters were estimated in a logistic regression model using, for energy and lipid intake, Q_{ik} measurements.

^bAll parameters were estimated in a logistic regression model using, for energy and lipid intake, predicted values computed in the calibration model.

^cfor a 5% increase

^dfor a 15% increase

^efor 5 years increase

^ffor 5 kg increase

^gfor 5 cm increase

Table 3: Validation of results using simulated data assuming conditional independence between $Q_{ik}|X_{ik}$ and $R_{ik}|X_{ik}$: Mean and associated 95% credible interval of estimates of true variance ($\hat{\Sigma}$), error variance ($\hat{\Sigma}_{\nu\epsilon}$), and log-odds ratio ($\hat{\beta}s$) obtained using actual data (model (a)) and simulated data (models (a) and (b)).

model	parameter	Actual data		Simulated data				
		Mean	model (a) 95% CI	Mean	model (a) 95% CI	Mean	model (b) 95% CI	
exposure	$\hat{\Sigma}_{11}^a$	0.155	(0.142, 0.170)	0.164	(0.152, 0.179)	0.147	(0.121, 0.176)	
	$\hat{\Sigma}_{12}^b$	0.165	(0.151, 0.182)	0.178	(0.163, 0.194)	0.163	(0.133, 0.196)	
	$\hat{\Sigma}_{22}^c$	0.247	(0.224, 0.273)	0.268	(0.244, 0.294)	0.251	(0.204, 0.311)	
	$\hat{\Sigma}_{\nu\epsilon 11}^d$	0.231	(0.198, 0.267)	0.229	(0.203, 0.256)	0.255	(0.215, 0.299)	
	$\hat{\Sigma}_{\nu\epsilon 12}^e$	—	—	—	—	0.023	(-0.008, 0.057)	
	$\hat{\Sigma}_{\nu\epsilon 22}^f$	1.029	(1.005, 1.053)	1.019	(0.997, 1.041)	1.035	(1.003, 1.067)	
	$\hat{\Sigma}_{\nu\epsilon 33}^g$	0.337	(0.281, 0.392)	0.341	(0.297, 0.391)	0.363	(0.295, 0.434)	
	$\hat{\Sigma}_{\nu\epsilon 34}^h$	—	—	—	—	0.014	(-0.048, 0.072)	
	$\hat{\Sigma}_{\nu\epsilon 44}^i$	2.081	(2.036, 2.127)	2.050	(2.008, 2.094)	2.061	(1.995, 2.124)	
	disease	$\hat{\beta}_2$	-0.043	(-0.141, 0.053)	0.017	(-0.080, 0.113)	0.021	(-0.085, 0.129)
		$\hat{\beta}_3$	0.431	(0.128, 0.751)	0.446	(0.327, 0.570)	0.449	(0.330, 0.576)
		$\hat{\beta}_4$	0.105	(-0.123, 0.339)	-0.022	(-0.240, 0.200)	0.027	(-0.259, 0.205)
$\hat{\beta}_5$		-0.098	(-0.644, 0.487)	-0.063	(-0.388, 0.272)	-0.072	(-0.404, 0.256)	

^atrue variance of energy intake

^btrue covariance between energy and lipid intake

^ctrue variance of lipid intake

^derror variance of energy intake in Q measurements

^eerror covariance of energy intake in Q and R measurements

^ferror variance of energy intake in R measurements

^gerror variance of lipid intake in Q measurements

^herror covariance of lipid intake in Q and R measurements

ⁱerror variance of lipid intake in R measurements

Table 4: Mean and associated 95% credible interval of each element of total ($\hat{\Sigma}$), between-center ($\hat{\Sigma}_x$) variance-covariance matrix ($\times 10$), and ICC values for true intakes in the EPIC exposure model, under the assumption of conditional independence between $Q_{ik}|X_{ik}$ and $R_{ik}|X_{ik}$.

$\hat{\Sigma}$	energy		lipids	
energy	0.155	(0.142, 0.170)	0.845 ^a	(0.836, 0.854)
lipids	0.165	(0.151, 0.182)	0.247	(0.224, 0.273)

$\hat{\Sigma}_x$	energy		lipids	
energy	0.077	(0.039, 0.151)	0.401 ^b	(-0.034, 0.728)
lipids	0.034	(-0.003, 0.091)	0.090	(0.045, 0.177)

	energy	lipids
ICC	0.331	0.267

^acorrelation coefficient between energy and lipids at the individual level

^bcorrelation coefficient between energy and lipids at the aggregate level

Table 5: Mean and associated 95% credible interval of each element of the error variance-covariance matrix ($\times 10$) for R_{ik} ($\hat{\Sigma}_\nu$) and Q_{ik} ($\hat{\Sigma}_\epsilon$) measurements in the EPIC exposure model, under the assumption of conditional independence between $Q_{ik}|X_{ik}$ and $R_{ik}|X_{ik}$.

$\hat{\Sigma}_\nu$	energy		lipids	
energy	1.029	(1.005, 1.053)	0.853 ^a	(0.849, 0.856)
lipids	1.248	(1.218, 1.278)	2.081	(2.036, 2.127)

$\hat{\Sigma}_\epsilon$	energy		lipids	
energy	0.231	(0.198, 0.267)	0.926 ^b	(0.912, 0.941)
lipids	0.258	(0.222, 0.297)	0.337	(0.281, 0.392)

^acorrelation coefficient between errors of energy and lipids in R measurements

^bcorrelation coefficient between errors of energy and lipids in Q measurements

Table 6: Mean and associated 95% credible interval for fixed and random effects in the EPIC measurement and disease models, under the assumption of conditional independence between $(Q_{ik}|X_{ik})$ and $(R_{ik}|X_{ik})$.

model	exposure	fixed effects			random effects		
			Mean	95% CI		Mean ^a	95% CI
measurement							
	energy	$\hat{\alpha}_{21}$	1.688	(1.522, 1.822)	$\hat{\sigma}_{\lambda_{21}}$	0.262	(0.183, 0.369)
	lipids	$\hat{\alpha}_{22}$	1.659	(1.486, 1.805)	$\hat{\sigma}_{\lambda_{22}}$	0.247	(0.172, 0.348)
disease							
	intercept	$\hat{\beta}_1$	-4.443	(-4.701, -4.213)	$\hat{\sigma}_{\gamma_1}$	0.385	(0.197, 0.632)
	energy ^b	$\hat{\beta}_2$	-0.043	(-0.141, 0.053)	$\hat{\sigma}_{\gamma_2}$	0.290	(0.118, 0.567)
	mean energy ^b	$\hat{\beta}_3$	0.431	(0.128, 0.751)			
	lipids ^c	$\hat{\beta}_4$	0.105	(-0.123, 0.339)	$\hat{\sigma}_{\gamma_3}$	0.316	(0.142, 0.574)
	mean lipids ^c	$\hat{\beta}_5$	-0.098	(-0.644, 0.487)			
	age ^d	$\hat{\beta}_6$	0.148	(0.068, 0.228)			
	mean age ^d	$\hat{\beta}_7$	0.464	(0.154, 0.777)			
	weight ^e	$\hat{\beta}_8$	0.027	(-0.025, 0.077)			
	mean weight ^e	$\hat{\beta}_9$	-0.180	(-0.567, 0.210)			
	height ^f	$\hat{\beta}_{10}$	0.121	(0.023, 0.220)			
	mean height ^f	$\hat{\beta}_{11}$	-0.055	(-0.584, 0.470)			

^aSquare root of variance components

^bfor 5% increase

^cfor 15% increase

^dfor 5 years increase

^efor 5 kg increase

^ffor 5 cm increase

Table 7: Mean and associated 95% credible interval of each element of total ($\hat{\Sigma}$), between-center ($\hat{\Sigma}_x$) variance-covariance matrix ($\times 10$) and ICC values for true intakes in the EPIC exposure model, assuming conditional dependence between $Q_{ik}|X_{ik}$ and $R_{ik}|X_{ik}$.

$\hat{\Sigma}$	energy		lipids	
energy	0.018	(0.012, 0.023)	0.604 ^a	(0.539, 0.664)
lipids	0.032	(0.024, 0.043)	0.164	(0.126, 0.206)

$\hat{\Sigma}_x$	energy		lipids	
energy	0.076	(0.038, 0.149)	0.395 ^b	(-0.040, 0.727)
lipids	0.033	(-0.003, 0.090)	0.089	(0.045, 0.175)

	energy	lipids
ICC	0.687	0.303

^acorrelation coefficient between energy and lipids at the individual level

^bcorrelation coefficient between energy and lipids at the aggregate level

Table 8: Mean and associated 95% credible interval of each element of the unified error variance-covariance matrix ($\times 10$) ($\hat{\Sigma}_{\nu\epsilon}$) for R_{ik} and Q_{ik} measurements in the EPIC exposure model, assuming conditional dependence between $Q_{ik}|X_{ik}$ and $R_{ik}|X_{ik}$.

Measurements	$\hat{\Sigma}_{\nu\epsilon}$	energy		lipids	
R	energy	1.178	(1.154, 1.201)	0.871 ^a	(0.865, 0.878)
	lipids	1.398	(1.368, 1.429)	2.186	(2.126, 2.244)

		energy		lipids	
Q	energy	0.529	(0.501, 0.553)	0.964 ^a	(0.949, 0.977)
	lipids	0.594	(0.559, 0.662)	0.716	(0.662, 0.766)

		energy		lipids	
R and Q		0.287 ^b	(0.276, 0.297)	0.173 ^b	(0.147, 0.197)

^acorrelation coefficient between errors in energy and lipids

^bcorrelation coefficient between errors in R and Q

Table 9: Mean and associated 95% credible interval for fixed and random effects in the EPIC measurement and disease models, assuming conditional dependence between $Q_{ik}|X_{ik}$ and $R_{ik}|X_{ik}$.

model	exposure	fixed effects			random effects		
			Mean	95% CI	Mean ^a	95% CI	
measurement							
	energy	$\hat{\alpha}_{21}$	2.560	(1.508, 3.145)	$\hat{\sigma}_{\lambda_{21}}$	1.018	(0.619, 1.602)
	lipids	$\hat{\alpha}_{22}$	1.214	(0.760, 1.538)	$\hat{\sigma}_{\lambda_{22}}$	0.353	(0.195, 0.476)
disease							
	intercept	$\hat{\beta}_1$	-4.450	(-4.717, -4.209)	$\hat{\sigma}_{\gamma_1}$	0.399	(0.201, 0.666)
	energy ^b	$\hat{\beta}_2$	0.075	(-0.185, 0.342)	$\hat{\sigma}_{\gamma_2}$	0.391	(0.125, 0.878)
	mean energy ^b	$\hat{\beta}_3$	0.450	(0.129, 0.801)			
	lipids ^c	$\hat{\beta}_4$	0.101	(-0.111, 0.316)	$\hat{\sigma}_{\gamma_3}$	0.338	(0.148, 0.620)
	mean lipids ^c	$\hat{\beta}_5$	-0.108	(-0.701, 0.494)			
	age ^d	$\hat{\beta}_6$	0.152	(0.072, 0.231)			
	mean age ^d	$\hat{\beta}_7$	0.443	(0.104, 0.780)			
	weight ^e	$\hat{\beta}_8$	0.026	(-0.026, 0.076)			
	mean weight ^e	$\hat{\beta}_9$	-0.142	(-0.545, 0.291)			
	height ^f	$\hat{\beta}_{10}$	0.117	(0.019, 0.216)			
	mean height ^f	$\hat{\beta}_{11}$	-0.084	(-0.661, 0.484)			

^aSquare root of variance components

^bfor 5% increase

^cfor 15% increase

^dfor 5 years increase

^efor 5 kg increase

^ffor 5 cm increase