American Journal of Epidemiology
Vol 138, No 6

A Bayesian Approach to Measurement Error Problems in Epidemiology Using Conditional Independence Models

Sylvia Richardson[1] and Walter R. Gilks[2]

Risk factors used in epidemiology are often measured with error which can seriously affect the assessment of the relation between risk factors and disease outcome. In this paper, a Bayesian perspective on measurement error problems in epidemiology is taken and it is shown how the information available in this setting can be structured in terms of conditional independence models. The modeling of common designs used in the presence of measurement error (validation group, repeated measures, ancillary data) is described. The authors indicate how Bayesian estimation can be carried out in these settings using Gibbs sampling, a sampling technique which is being increasingly referred to in statistical and biomedical applications. The method is illustrated by analyzing a design with two measuring instruments and no validation group. *Am J Epidemiol* 1993;138:430–42

biometry; Bayesian method; epidemiologic methods; Monte Carlo method

It is widely recognized that risk factors (exposures or more generally covariates) used in epidemiology are often measured with error which can seriously affect the assessment of the relation between risk factors and disease outcome. In some instances, it is possible to seek to improve the measuring instrument and thus have a better record of exposure for the whole population. There are many situations, however, where it is not feasible to obtain accurate measurements on the entire study population, although this might be attempted on a smaller subset (the *validation group*). Examples of this type of design abound in the field of nutritional or occupational epidemiology. It is thus important, at the analysis stage, to have statistical methods which can successfully estimate the strength of the association between the risk factor and the disease outcome by using some (often very partial) information on the measurement instrument.

It is clear that any method proposed for correcting parameter estimates in the presence of measurement error is strongly dependent on some knowledge of the measurement error process. How best to integrate this knowledge in the context of epidemiologic studies has been the subject of much interest (1–6). In the frequentist (non-Bayesian) inference framework, the methods proposed for correcting relative risk estimates meet with certain difficulties which have been solved in a variety of ways. Special features of the model can be exploited, for example, considering nondifferential symmetric misclassification errors in case-control studies (2). Alternatively, conditions allowing approximate inference might be imposed either on the disease process (the rare disease assumption (3)), or on the measurement error process (small error-variance assumptions (7)).

In this paper, we aim to give a Bayesian

Received for publication December 2, 1992, and in final form May 26, 1993.

Abbreviation. RR, relative risk.

[1] Unité 170, Institut National de la Santé et de la Recherche Médicale, Villejuif, France
[2] Medical Research Council, Biostatistics Unit, Cambridge, England
Reprint requests to Dr. S. Richardson, Unité 170, INSERM, 16 Avenue Paul Vaillant-Couturier, 94807 Villejuif Cedex, France.

perspective on measurement error problems and to show how, with this new perspective, many of the difficulties encountered by existing methods can be overcome. Bayesian methods have not been widely used in epidemiology even though a large range of problems can be formulated naturally in this framework (8). Recent statistical and computational advances have now made this approach possible in many applications, in particular in biomedical research (9, 10). Recently, Thomas et al. (11) have used a Bayesian approach to analyze a particular case of measurement error in matched case-control studies.

BAYESIAN APPROACH

General principles

Bayesians and frequentists differ through their concept of probability. As a consequence, they handle uncertainty in model parameters differently. Bayesians think of model parameters as random variables, and they interpret the probability distribution of a model parameter in terms of degrees of belief about values of that parameter. Frequentists think of probabilities as frequencies observed in a long run of repeated experiments, and they view model parameters as fixed (non-random) quantities which therefore cannot have probability distributions.

In the Bayesian approach, information available at the start of the study leads to specification of the *prior distribution* of the parameters. Once data have been gathered, inference is made on the basis of the *posterior distribution* of the parameters given the data, which, by Bayes theorem, are proportional to the product of the likelihood and the prior distribution. From this posterior distribution, point and interval estimates of the parameters might be computed.

An illustration of prior to posterior updating is given in figure 1. Let us consider, say, a relative risk (denoted by RR). From its prior distribution (top of figure 1), we can see that before any data were gathered, it was thought that the true value of the relative

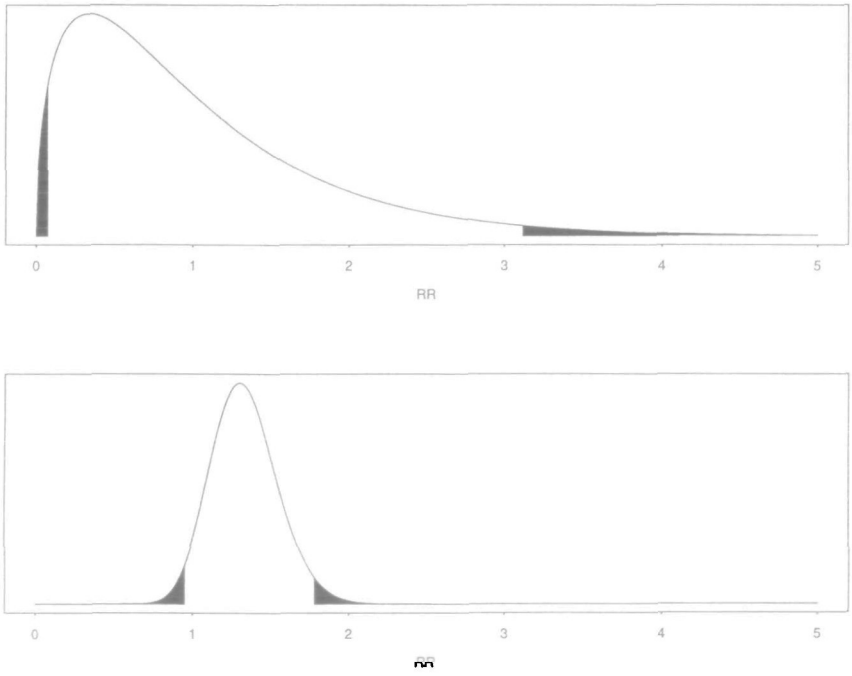

FIGURE 1. Illustration of prior to posterior updating in Bayesian inference. Top, prior distribution of RR; bottom, posterior distribution of RR.

risk lay between 0.07 and 3.12 with probability 0.95. The posterior distribution (bottom of figure 1) is more concentrated than the prior distribution, showing a probability of 0.95 that the relative risk lies between 0.95 and 1.78. This is called a 95 percent credibility interval. Note that a credibility interval has a natural interpretation in terms of probabilities. The frequentist analogue of a credibility interval is a confidence interval, which has a more difficult probabilistic interpretation: 95 percent of 95 percent confidence intervals for RR calculated in a long run of repeated experiments would contain the true value of RR. Not surprisingly, confidence intervals are often erroneously interpreted.

Measurement error problems from a Bayesian perspective

In the measurement error problem, some risk factors X related to the disease status Y are unknown, although surrogate measures Z of X are recorded. Estimation of the epidemiologic parameters of interest, e.g., the relative risks linking X and Y, must take into account all the uncertainty on X.

At the start of the study, some information (possibly limited) is available on the distribution of the risk factors in the general population, leading to the specification of a prior distribution for X. If the study is one in a series carried out on the same population, this prior information might be well focused, otherwise the prior distribution will be assumed to be suitably vague. During the study, surrogate measures are recorded which provide information on X. Therefore, having the surrogate measures reduces the uncertainty on the unknown risk factors X. A similar process operates for the epidemiologic parameters. From a literature review, some information might be gathered which allows a prior distribution to be specified for these parameters. The prior distribution would, for example, exclude unrealistically high or low relative risks. In the study data, information on the relative risks is contained in the disease status Y, but this information is strengthened by the surrogates Z through their link with the unknown risk factors X.

In the following sections, we shall show how we can structure the information available in this general epidemiologic setting of measurement error in terms of *conditional independence models*, and outline a method of estimation called Gibbs sampling.

CONDITIONAL INDEPENDENCE MODELING AND ITS ASSOCIATED GRAPHICAL REPRESENTATION

The construction of conditional independence models is carried out in two stages. The structural part of the model is set up first. This is followed by the specification of the functional part where all the distributions involved are precisely defined. In this section, we concentrate on the formulation of the structural part of measurement error problems.

A key concept in building the structure of a model is that of conditional independence. To illustrate this concept in our context, let us consider the three variables Z_i, X_i, and Y_i for an individual i. To state that *Y_i and Z_i are conditionally independent given X_i* is equivalent to making the classical assumption that the surrogate measures Z_i do not provide any information on the disease status Y_i if true values of risk factors X_i are known.

Using conditional independence, each variable in the model will be related conditionally to only a *few* other variables. Let us stress, at this stage, that the conditional independence assumptions contribute information on the structure of the model which will strengthen inference, presuming of course that these assumptions are correct. Therefore, great care has to be given to the consequences of each of the conditional independence assumptions. In this way, complex problems are broken down into modular components which have a relatively simple structure. In the measurement error problem, following the terminology intro-

duced by Clayton (12), two of the components are: a disease model, which expresses the relation between unknown risk factors X, possibly also some known risk factors C and the disease status Y; and a measurement model, which expresses the relation between the surrogate measures Z and the true unknown risk factors X.

Measurement error has traditionally been modeled in two different ways. In the classical measurement error formulation, the conditional distribution of the surrogates Z given the true risk factors X is specified, while in the Berkson formulation, it is the conditional distribution of X given Z which is specified.

Graphs

Conditional independence models are naturally expressed by means of a graphical representation (an influence diagram or graph) in which nodes represent random variables of interest and edges (arrows) reflect local dependencies.

The basic features of the graph corresponding to the classical measurement error situation are represented in the top part of figure 2. In this graph, square nodes denote known quantities (data), and circular nodes denote unknown quantities (unknown risk factors). The arrows entering node Y show that disease status depends on the risk factors X and C. The arrow entering Z reflects the classical measurement error hypothesis, i.e., that the distribution of Z is specified conditionally on X. The absence of an edge between Y and Z shows that they are *conditionally independent given* X, a classical assumption (see above).

The purpose of the top part of figure 2 is to convey the essential relations between the three groups of variables: the disease status, the risk factors, and the surrogates. Thus, each node stands generically for a family of variables. For example, X represents one or several risk factors while Z stands for all the surrogate measures of X, including repeats.

The parameters linking the different groups of variables need now to be specified. These parameters form three groups: 1) the epidemiologic parameters β (e.g., relative risks) modeling the link between the risk factors X and C and the disease status Y; 2) the measurement error parameters λ (e.g., measurement error variance) modeling the link between X and its surrogate Z; and 3) the exposure parameters π (e.g., population mean and variance) modeling the population distribution of X. Note that we are now distinguishing a third component, the exposure model, which describes the distribution of the unknown risk factors in the general population, i.e., the population distribution of X.

The structure of the graph represented in the top part of figure 2 can be enriched by the inclusion of these parameters (middle part of figure 2). Now two arrows, originating from X and λ, are pointing toward Z. By this, we indicate that the conditional distribution of Z is entirely specified given X and λ and similarly for the conditional distribution of Y given X, C, and β and that of X given π.

Model conditionals

The structure of the measurement error problem can thus be formalized by writing three equations, each expressing local dependencies:

$$\text{disease model} \quad [Y_i \mid X_i, C_i, \beta] \tag{1}$$

$$\text{measurement model} \quad [Z_i \mid X_i, \lambda] \tag{2}$$

$$\text{exposure model} \quad [X_i \mid C_i, \pi], \tag{3}$$

where the index i denotes individual i, and $[U \mid V]$ generically denotes the conditional

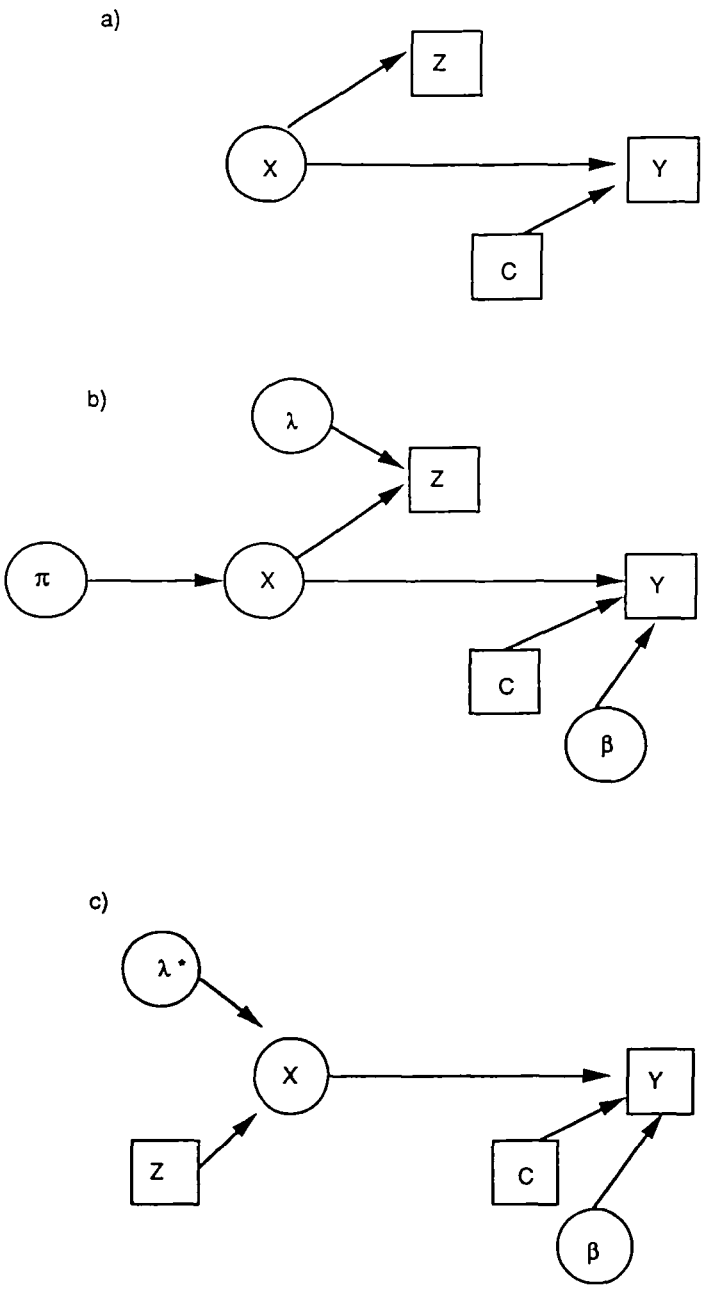

FIGURE 2. Measurement error graphs. a) conditional relations in a classical measurement error situation between the disease status Y, the risks factors X (unknown) and C (known) and the surrogate measure Z of X; b) graph corresponding to the classical measurement error situation modeled in equations 1, 2, and 3, and c) graph corresponding to the Berkson error model defined by equations 1 and 2′.

distribution of U given V. Since we are in a Bayesian framework, prior distributions for β, λ, and π are also required (denoted respectively by $[\beta]$, $[\lambda]$, and $[\pi]$). The description of the structure is completed by specifying that the joint distribution of all the variables can

be written as the product of all the model conditionals:

$$[\beta][\lambda][\pi] \prod_i [X_i \mid C_i, \pi] \prod_i [Z_i \mid X_i, \lambda] \prod_i [Y_i \mid X_i, C_i, \beta]. \quad (4)$$

Equation 4 can be broadly translated by saying that there are no local dependencies other than those stated by the model equations 1, 2, and 3; equivalently, equation 4 implies many additional conditional independence assumptions. For example, we have not included in equation 1 a dependence of Y_i on the risk factors $X_{i'}$ for other individuals i'. Thus, we assert that Y_i is independent of all the $X_{i'}$, $i' \neq i$, conditionally on X_i, C_i, and β. Similarly, equation 2 states that by conditioning on appropriately defined parameters λ and the true exposure X_i, the surrogate measures Z_i are independent among individuals. The construction of λ requires careful attention and will be discussed later.

Model equation 2 corresponds to a classical error model formulation. For a Berkson type of error model, equation 2 would be replaced by equation 2′,

$$\text{measurement model} \quad [X_i \mid Z_i, \lambda^*], \quad (2')$$

and there would be no need to specify an exposure model equation 3 since the distribution of X would be entirely specified by that of the known surrogates in this case (hence its attraction in the frequentist approach). This is a very strong assumption which has been debated at length (1, 13). In a Berkson error model, arrows would point from Z to X (bottom part of figure 2). This underlines the essential distinction between the classical and the Berkson error models and gives some intuition as to why the Berkson case is simpler to analyze.

INFERENCE

We now present some basic ideas concerning Bayesian inference in these models. Before carrying out the estimation, the functional part of the model has to be set up, which entails specifying explicitly the parametric form of all the conditional distributions. In some cases, there is a natural choice of parametrization, such as specifying a logistic distribution for the disease model. In other cases, the parametric distribution can be tailored to the particular application using some external information like the measurement and exposure model for atomic bomb survivors used by Pierce et al. (14). It will be important to undertake a sensitivity analysis when there is little or conflicting information to motivate the choice of parametrization.

Bayesian estimation of parameters is based on the posterior distribution of the parameters given the data. We use the word "parameter" here and below to mean both model parameter, β, λ, π, and unobserved data $\{X_i\}$. In our case, interest is ultimately in the epidemiologic parameters β. Nevertheless, there are many other unknown quantities: the other model parameters, λ and π, as well as the unknown risk factors X_i for each individual. Thus, our interest is really in the *marginal* posterior distribution of β given the data, i.e., the distribution of β when we don't know the values of the other parameters. Computing this marginal posterior distribution leads to a very high dimensional integral (since, for example, one would have to integrate over all the X_i) which is totally intractable. Estimation thus proceeds along a different line and makes use of a unifying computational method, Gibbs sampling (9, 15), which generates samples from the *joint posterior distribution* of all parameters (β, λ, π, and $\{X_i\}$) given the data. (The joint posterior distribution is proportional to the expression 4). From these samples, inference can be made straightforwardly on the parameters of interest, singly, jointly, or for any functions of parameters.

It is beyond the scope of this paper to describe this algorithm in depth and details of its

application for analyzing some measurement error problems can be found in Gilks and Richardson (16) and Richardson and Gilks (17). We restrict ourselves to describing mainly one essential component, the updating cycle. At first, arbitrary starting values for each parameter are chosen. Then, in turn, one parameter at a time is updated by sampling a new value for that parameter from its conditional distribution (or density) given the data and the current values of all other parameters in the model, referred to as the *current full conditional distribution*.

A cycle of the Gibbs sampler is completed when all the unknown variables in the model have been updated once. The updating cycle is repeated a large number of times. It has been shown that this process converges toward the distribution of interest, i.e., that the samples generated can be considered after a while as samples from the joint posterior distribution of all the parameters.

Clearly, a key step in the implementation of this algorithm is the derivation of the full conditional distribution for each parameter. This is simply proportional to the product of all terms which contain that parameter in the joint distribution given in equation 4. The full conditional distribution for the measurement parameters in a particular design will be given in the next section.

DEVELOPING A BAYESIAN GRAPHICAL MODELING APPROACH FOR REPRESENTING EPIDEMIOLOGIC DESIGNS

To give relative risk estimates in the presence of measurement errors, different epidemiologic designs have been proposed, e.g., the use of a validation group, the inclusion of several measuring instruments, and/or the inclusion of repeated measures.

Design with a validation group

In this design, the existence of a gold standard (an error-free method for measuring risk-factor X) is assumed. This gold standard is usually only available on a small subset of the population, called the validation group. Both X and its surrogate measures Z are recorded in the validation group, while in the main study only the surrogate measures Z will be available for each individual. Nevertheless, information about λ has somehow to be transferred from the validation group to the main study. In other words, from the comparison of X and Z in the validation group, information on the unknown risk factors X for the main study has to be gained. When knowledge of the disease status is also known for the individuals in the validation group, this group in referred to as internal; otherwise, the validation group is called external.

Figure 3 gives a graphical representation of a design with an internal validation group. The structure of the two parts of the graph is similar except that X is drawn as a square box in the validation group (top of the graph) since X is presumed known in this group. The graph shows how the validation group contributes information on λ, therefore providing information on the relation between X and Z in the main study, and thereby strengthening inference about β.

To detail more fully one step in the implementation of the Gibbs sampling algorithm, let us write the full conditional distribution for the measurement error parameters in this design. It is proportional to

$$[\lambda] \cdot \prod_{i \in P_1} [Z_i \mid X_i, \lambda] \cdot \prod_{i \in P_2} [Z_i \mid X_i, \lambda],$$

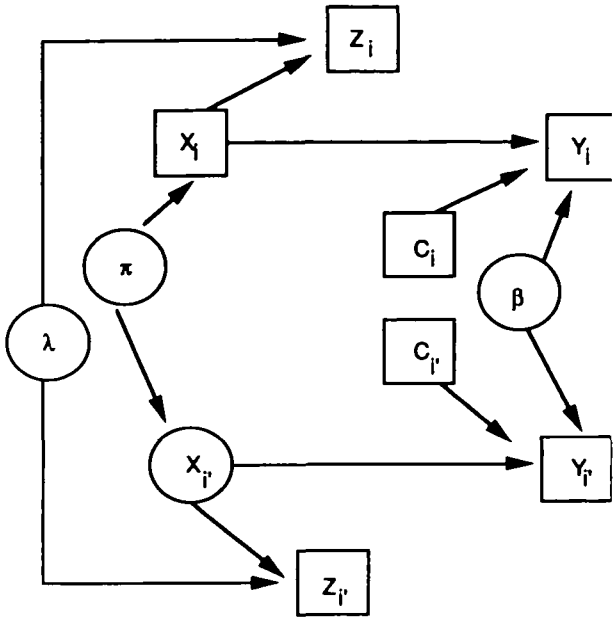

FIGURE 3. Graph corresponding to a design with an internal validation group and a main study.

where P_1 denotes the individuals included in the validation group, P_2 those in the main study and the symbol $i \epsilon P_t$ indicates that the product is taken over all the individuals in P_t, $t = 1, 2$. Thus, the parameters λ are updated using the information in the validation group (which will not vary) and simulated risk factors X_t (which will change at each iteration) for the main study individuals.

In the algorithm we are sampling from the joint posterior distribution of the parameters (models parameters and unobserved risk factors), the transfer of information between different parts of the design is effected appropriately. Hence, the full uncertainty in the true risk factors X_t is taken into account in the estimation of the epidemiologic parameters β.

Design with several measuring instruments and repeated measures

In some situations, it is unrealistic to assume the existence of a gold standard. Nevertheless, some information on the measuring process can sometimes be gained by using repeated measures and/or by combining several instruments. Let us first outline the general structure of such designs, a structure which is discussed in detail in Richardson and Gilks (17).

Denote by Z_{imr} the result of the rth repeated measure by measuring instrument m of the true risk factor X_i. We now consider situations where, conditionally on X_i and on measurement parameters λ_m, it is reasonable to suppose that the Z_{imr} are independent between repeats and also between instruments.

The measurement model equation 2 is thus replaced by

$$[Z_{imr} \mid X_i, \lambda_m]. \tag{5}$$

The assumed conditional independence relation between repeats and instruments imply that in expression 4 of the joint distribution and hence that of the posterior distribution:

$$\prod_i [Z_i \mid X_i, \lambda] \quad \text{is replaced by} \quad \prod_{i,m,r} [Z_{imr} \mid X_i, \lambda_m].$$

Note that the measurement model formulated in equation 5 is fairly general but would not fit all measurement processes, for example, designs where the same measuring instrument is used for several covariates. An appropriate modification of the structure of the measurement model equation could easily be defined to accommodate this situation, but, for the sake of clarity, we will stay with the setting defined by equation 5. Let us now detail the functional part of the measurement model 5 for a particular design. The functional part of the exposure and disease models will be detailed in the next section.

An example. We suppose that there are two measuring instruments, the first one having low precision but being unbiased and relatively cheap to administer, while, by contrast, the second measuring instrument has a higher precision but is costly and known to be biased. To be precise, we suppose that the conditional distribution of the first unbiased instrument is given by a Normal distribution with mean X_i and variance θ_1^{-1} denoted by

$$[Z_{i1r} \mid X_i, \theta_1] \sim \mathrm{N}(X_i, \theta_1^{-1}), \tag{6}$$

and so the only measurement parameter in λ_1 is θ_1, the precision (inverse of the variance) of Instrument 1. For the second biased instrument, the conditional distribution is also a Normal distribution:

$$[Z_{i2r} \mid X_i, \phi_2, \psi_2, \theta_2] \sim \mathrm{N}(\phi_2 + \psi_2 X_i, \theta_2^{-1}), \tag{7}$$

and the measurement error parameters λ_2 are the intercept and slope parameters ϕ_2 and ψ_2, expressing the linear relation between the true exposure and its surrogate, and the precision θ_2.

Repeats of an instrument only provide information on its precision. Thus, in general, the parameters ϕ_2 and ψ_2 cannot be sensibly estimated without the knowledge of a gold standard measured in a validation subgroup. We now suppose that we have set up a design where, for a subgroup of individuals, Instrument 1 has been repeated twice and Instrument 2 has also been recorded. In this particular design, even though there is no gold standard, the data

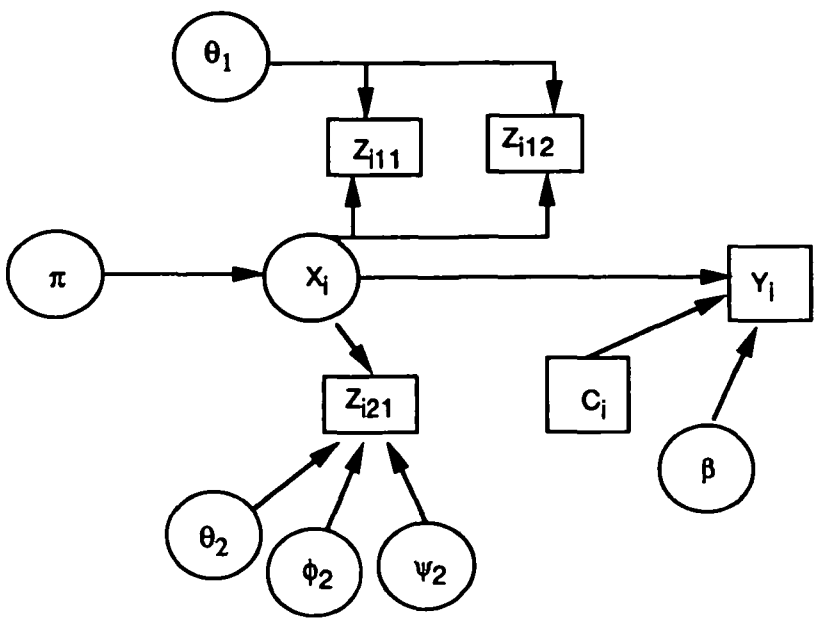

FIGURE 4. Graph corresponding to a design with two measuring instruments and no validation group as defined in equations 6 and 7

still contain information on ϕ_2 ψ_2 because there is information on X_i from the repeats of the unbiased Instrument 1 (see figure 4), information which is used as a "simulated gold standard" in order to estimate ϕ_2 and ψ_2.

The corrections of relative risk estimates offered by designs of this kind will be illustrated in the Results section.

Design with ancillary risk factor information

In occupational or environmental epidemiology, risk factor information for each individual is often not directly available but has to be constructed using external information and a group-level characteristic of the individual. A typical example is that of the use of job-exposure matrices in industrial epidemiology. Job-exposure matrices provide information on exposures to each of many industrial agents in each of many finely subdivided categories of occupation. They are commonly constructed by industrial experts from detailed job descriptions obtained in a specially conducted survey. In some study designs, the exposure of an individual to industrial agents is then characterized using only his job title and ancillary risk factor information contained in a job-exposure matrix. The measurement error model implied by this design is different from those considered previously as imprecisions in exposure information provided by the survey need to be taken into account in a survey model which is then linked to the exposure model.

Denoting by π_{jk} the underlying (unobserved) probability of being exposed to agent k (dichotomous exposure) in job j, we can model the conditional distribution of m_{jk} the observed number of people in the survey with job j found exposed to agent k, as a binomial distribution:

$$[m_{jk} \mid \pi_{jk}, n_j] = \text{Binomial} \quad (\pi_{jk}, n_j),$$

where n_j is the number of people with job j included in the survey.

This survey model is then linked to the exposure X_{ik} of individual i to agent k through the job title $j = j(i)$ of individual i by assuming that X_{ik} is exposed with probability $\pi_{j(i)k}$:

$$[X_{ik} \mid \pi_{j(i)k}] = \text{Bernoulli} \quad (\pi_{j(i)k}).$$

The disease model is unchanged (see figure 5 for the graph associated with this model). Note that the survey does not provide direct information on X but rather on the prior distribution

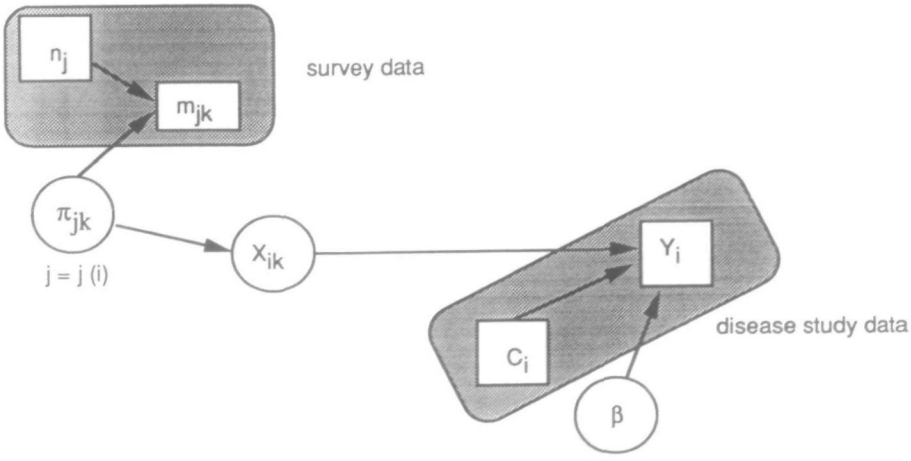

FIGURE 5. Graph corresponding to a design with ancillary risk factor information coming from a survey

TABLE 1. Gibbs sampling analysis of the simulated data set: means ± standard deviation (SD)

Parameter	True value	Gibbs sampling analysis Posterior			Analysis on true covariates Posterior
		Mean ± SD	2 5%	97 5%	Mean ± SD
Baseline risk*	0 45	0.36 ± 0 08	0.22	0 50	0 36 ± 0.04
RR_X†	2 45	3 62 ± 1 21	2 27	6.41	3 17 ± 0 34
RR_C‡	3.32	3 52 ± 0.46	2 68	4.47	3 38 ± 0.37
θ_1	0 9	0 93 ± 0.09	0 78	1.12	
ϕ_2	0 8	0.80 ± 0 04	0 71	0.88	
Ψ_2	0.4	0 37 ± 0.05	0 29	0 47	
θ_2	5 0	4.39 ± 0 53	3 43	5.52	

* The baseline risk corresponds to e^{β_0}
† $RR_X = e^{\beta_1}$
‡ $RR_C = e^{\beta_2}$

of X. We are thus neither in a classical measurement error situation nor in a Berksonian one. (See Gilks and Richardson (16) for details showing the good performance of Bayesian modeling for analyzing designs of this kind.)

ANALYSIS OF A SIMULATED DATA SET

To illustrate the performance of our Bayesian estimation approach, we present here the analysis of a simulated data set reproducing a design where the measurement parameters are apprehended through the combination of two measuring instruments, as in the second example of the previous section.

Design set-up

Two risk factors are involved in the disease model. The first risk factor, X, is measured with error and the second risk factor, C, is known accurately. We consider the case of a logistic link between risk factors and disease status. To be precise, we suppose that Y_i follows a Bernoulli distribution with parameter α_i, where logit $\alpha_i = \beta_0 + \beta_1 X_i + \beta_2 C_i$. We suppose that the exposure vector (X, C) follows a bivariate normal distribution, with mean μ and variance-covariance matrix Σ.

The study is designed to include two parts (i.e., two subgroups of individuals) which differ only with respect to their measurement process. There are 200 individuals in Part 1 and 1,000 individuals in Part 2. In Part 1, the measurement model follows exactly that described in the example given previously, that is Instrument 1 has been measured twice (equation 6) and Instrument 2 has been recorded once (equation 7) for all 200 individuals. In Part 2, only Instrument 1 has been recorded for the 1,000 individuals (equation 6).

A data set was generated using "true" values of $\beta = (\beta_0, \beta_1, \beta_2)$, θ_1, ϕ_2, ψ_2, and θ_2 given in table 1 (column: "true values") and with

$$\mu = \begin{pmatrix} 0.5 \\ -0.5 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} 1.02 & 0.56 \\ 0.55 & 0.96 \end{pmatrix}.$$

Thus, we have simulated a situation with two detrimental risk factors X and C corresponding to relative risks: $RR_X = 2.46$ and $RR_C = 3.32$, respectively, with a positive correlation

between X and C ($r = 0.56$). Note that Instrument 2 is substantially more accurate than Instrument 1 ($\theta_2 > 5\theta_1$).

RESULTS

Table 1 presents the results from the Gibbs sampling analysis of the simulated data set. We have summarized the marginal posterior distribution of the parameters of interest by reporting their means and standard deviations, and 2.5 and 97.5 percentiles. In the last column of table 1, we have given, as benchmarks, the estimates of the relative risks obtained by a Bayesian logistic regression analysis based on the knowledge of the true values of X_i for the 1,200 individuals in the study. Note that the estimate of RR_X (3.17), which would be obtained in the absence of measurement error in our data set, is a little higher than its "true value" (2.45) used to produce the simulated data set.

The results show that our estimation method has performed very satisfactorily with all the estimated values lying well within one posterior standard deviation of the values given by the analysis on "true covariates" (last column). As expected, the posterior standard deviation for RR_C which corresponds to the relative risk for the covariate C measured without error is smaller than that for RR_X. We note some degree of skewedness in the distribution of the estimate of the relative risk RR_X. It is further interesting to check that the measurement parameters for Instrument 2 have been well estimated, even though our design did not include a validation subgroup. This highlights how information has been naturally propagated between the two measuring instruments and between the two parts of the design. This is a key feature of our estimation method not shared by other methods.

DISCUSSION

There is a large literature on methods for analyzing epidemiologic data with measurement errors. In this paper, we have presented a Bayesian approach, based on conditional independence modeling, where estimation is carried out by simulations (Gibbs sampling). A Gibbs sampling approach has also been used by Thomas et al. (11) in their analysis of the leukemia risk in Utah following radioactive fallout.

Some key features of this approach make it particularly fruitful in the epidemiologic context. First, the method is flexible and adapts to the structure of the error problem without making artificial assumptions. We have illustrated this point by discussing a few specific designs, but we must stress that many complex measurement error problems can be tackled in this way. Second, this approach gives a framework in which different sources of information can be integrated. Our example was concerned with combining two different measuring instruments; another example which could be modeled straightforwardly is that of combining ancillary information from a job-exposure matrix and individual exposure assessment by experts in industrial epidemiology. It is also easy to see that missing data are simple to handle as they can be treated as additional unknown parameters.

The approach that we have developed is fully parametric. As for other methods dealing with measurement errors, misspecification of the conditional distributions involved in the measurement or the exposure models will perturb the inference on the epidemiologic parameters of interest. Interesting lines for further research are related to how one can detect misspecification and what is the sensitivity of our method to a wrong choice. Some authors have investigated nonparametric approaches (18, 19) for estimating the measurement model, and Thomas et al. (11) have combined a nonparametric specification of the exposure model with Gibbs sampling. There is scope for further work in this direction.

Another important advantage of our approach is that all sources of errors are considered and the resulting uncertainty is taken fully into account in the estimation of the parameters of interest. In particular, imprecision in the parameters λ of the measurement model are fully allowed for in calculating the variability of the regression parameters β. In contrast, substitution methods (say from a validation subgroup to a main study (3)) only make approximate allowance for uncertainty of estimation in the validation subgroup.

The method that we have outlined will allow designs of substantial complexity to be analyzed. A challenging issue is now to use this powerful tool for the planning and design of epidemiologic studies.

REFERENCES

1. Armstrong BG. The effects of measurement errors on relative risk regression. Am J Epidemiol 1990;132:1176–84.
2. Duffy SW, Maximovitch DM, Day NE. External validation, repeat determination and precision of risk estimation in misclassified exposure data in epidemiology. J Epidemiol Community Health 1992;46:620–4.
3. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for measurement error: the case of multiple covariate measured with error. Am J Epidemiol 1990;132:734–45.
4. Rosner B, Spiegelman D, Willett WC. Correction of logistic regression relative risk estimates and confidence intervals for random within-person measurement error. Am J Epidemiol 1992; 136:1400–13.
5. Thomas D, Stram D, Dwyer J. Exposure measurement error: influence on exposure-disease relationship and methods of correction. Ann Rev Public Health, 1993;14 (in press).
6. Willett W. An overview of issues related to the correction of non-differential exposure measurement error in epidemiologic studies. Stat Med 1989;8:1031–40.
7. Whittemore AS, Keller JB. Approximations for regression with covariate measurement error. J Am Stat Assoc 1988;83:1057–66.
8. Breslow N. Biostatistics and Bayes. Stat Sci 1990;5:269–98.
9. Gelfand AE, Hills SE, Racine-Poon A, et al Illustration of Bayesian inference in normal data models using Gibbs sampling. J Am Stat Assoc 1990;85:972–85.
10. Gilks WR, Clayton DG, Spiegelhalter DJ, et al. Modeling complexity: applications of Gibbs sampling in medicine (with discussion). J R Stat Soc [B] 1993;55:39–52.
11. Thomas DC, Gauderman J, Kerber R. A nonparametric Monte Carlo approach to adjustment for covariate measurement errors in regression analysis. Biometrics (in press).
12. Clayton DG. Models for the analysis of cohort and case-control studies with inaccurately measured exposures. In: Dwyer JH, Feinleib M, Lippert P, et al, eds. Statistical models for longitudinal studies of health. New York: Oxford University Press, 1991:301–31
13. Caroll RJ. Covariance analysis in generalized linear measurement error models. Stat Med 1989;8 1075–93.
14. Pierce DA, Stram DO, Vaeth M, et al. The errors in variables problem: considerations provided by radiation dose-response analyses of the A-bomb survivor data. J Am Stat Assoc 1992;87:351–59.
15. Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. Institute of Electrical and Electronic Engineers Transactions in Pattern Analysis and Machine Intelligence 1984,6:721–41.
16. Gilks W, Richardson S. Analysis of disease risks using ancillary risk factors, with application to job-exposure matrices. Stat Med 1992;11:1443–63.
17. Richardson S, Gilks W. Conditional independence models for epidemiological studies with covariate measurement error. Stat Med (in press).
18. Pepe MS, Fleming TR. A non parametric method for dealing with mismeasured covariate data. J Am Stat Assoc 1991;86:108–13.
19. Caroll RJ, Wand MP. A semiparametric estimation in logistic measurement error models. J R Stat Soc [B] 1991;53:573–85.