A Monte Carlo Maximum Likelihood Method for Estimating Uncertainty Arising from Shared Errors in Exposures in Epidemiological Studies of Nuclear Workers

Author(s): Leslie Stayner, Martine Vrijheid, Elisabeth Cardis, Daniel O. Stram, Isabelle Deltour, Stephen J. Gilbert, and Geoffrey Howe
Source: Radiation Research, 168(6):757-763.
Published By: Radiation Research Society
https://doi.org/10.1667/RR0677.1
URL: http://www.bioone.org/doi/full/10.1667/RR0677.1

RADIATION RESEARCH **168**, 757–763 (2007)
0033-7587/08 $15.00

A Monte Carlo Maximum Likelihood Method for Estimating Uncertainty Arising from Shared Errors in Exposures in Epidemiological Studies of Nuclear Workers

Leslie Stayner,[a,1] Martine Vrijheid,[b] Elisabeth Cardis,[b] Daniel O. Stram,[c] Isabelle Deltour,[b] Stephen J. Gilbert[d] and Geoffrey Howe[e,2]

[a] *Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois, Chicago, Illinois 60612-4394;* [b] *Radiation Group, International Agency for Research on Cancer, 69 372 Lyon Cedex 08, France;* [c] *Division of Biostatistics and Genetic Epidemiology, Department of Preventive Medicine, University of Southern California, Los Angeles, California 90033;* [d] *Risk Evaluation Branch, Education and Information Division, National Institute of Occupational Safety and Health, Cincinnati, Ohio; and* [e] *Department of Epidemiology, Columbia University School of Public Health, New York, New York*

Stayner, L., Vrijheid, M., Cardis, E., Stram, D. O., Deltour, I., Gilbert, S. J. and Howe, G. A Monte Carlo Maximum Likelihood Method for Estimating Uncertainty Arising from Shared Errors in Exposures in Epidemiological Studies of Nuclear Workers. *Radiat. Res.* 168, 757–763 (2007).

Errors in the estimation of exposures or doses are a major source of uncertainty in epidemiological studies of cancer among nuclear workers. This paper presents a Monte Carlo maximum likelihood method that can be used for estimating a confidence interval that reflects both statistical sampling error and uncertainty in the measurement of exposures. The method is illustrated by application to an analysis of all cancer (excluding leukemia) mortality in a study of nuclear workers at the Oak Ridge National Laboratory (ORNL). Monte Carlo methods were used to generate 10,000 data sets with a simulated corrected dose estimate for each member of the cohort based on the estimated distribution of errors in doses. A Cox proportional hazards model was applied to each of these simulated data sets. A partial likelihood, averaged over all of the simulations, was generated; the central risk estimate and confidence interval were estimated from this partial likelihood. The conventional unsimulated analysis of the ORNL study yielded an excess relative risk (ERR) of 5.38 per Sv (90% confidence interval 0.54–12.58). The Monte Carlo maximum likelihood method yielded a slightly lower ERR (4.82 per Sv) and wider confidence interval (0.41–13.31).

INTRODUCTION

In occupational and environmental epidemiological studies, including those concerning radiation exposures, an exposure matrix or a physical dosimetry system is often used to estimate exposures (or doses) for individual subjects, and these estimates are then used in exposure–response analyses. However, the effect of errors in the estimated exposures on the exposure–response relationships has generally been ignored or at best discussed qualitatively.

Errors in elements of the exposure matrix or dosimetry system are a common concern in occupational and environmental studies. These errors result from the use of group means to impute individual exposures (or doses). Individuals sharing the same job in the same facility, or individuals in a similar location, in the same period are frequently assigned the same exposure estimates based on an ''exposure matrix'', and these exposure estimates represent the mean values of individual exposure for each period and plant/location. Shared errors may also occur, as in the example used in this paper (a study of nuclear workers), when individual estimates of dose are available and corrections are made to the dose estimates that are applied to all individuals within the same job or facility. If these mean doses (or exposures) are known without error, the random variation of individual doses around the true means produces what is known as ''Berkson'' error, which produces unbiased estimates of the dose–response relationship when the model is linear (*1*). Errors in the assignment of the means for each period introduce a new aspect of the problem, namely the sharing of errors between subjects. A primary effect of the sharing of errors due to misassignment of the mean values for each period is that the variances of the parameters in the dose–response relationships will be underestimated (*2*). This is true even when, as in fitting a linear dose–response relationship, risk estimates are unbiased when the dosimetry system contains only Berkson errors.

Relatively little work has been done on the development of methods to estimate the effect of such shared errors on the uncertainty of exposure–response relationships. The purpose of this paper is to illustrate the application of Monte Carlo maximum likelihood methods to the estimation of

[1] Address for correspondence: Division of Epidemiology and Biostatistics (MC 923), School of Public Health, University of Illinois, 1603 West Taylor Street, Chicago, IL 60612-4394; e-mail: lstayner@uic.edu.
[2] Deceased.

confidence intervals that reflect both ordinary statistical sampling error and error related to the uncertainty in the exposure estimates. The method is illustrated by application to an analysis of all cancer mortality risk related to radiation dose in a study of nuclear workers at the Oak Ridge National Laboratory (ORNL).

MATERIALS AND METHODS

Conceptual Model

We take as our starting point a Berkson model for measurement error in which there are both shared and unshared measurement errors. When dealing with dosimetry systems that may incorporate either biased errors or errors with a classical error structure, it is often helpful in the statistical analysis to attempt to transform the measurement error model into a Berkson one. For example, Thomas *et al.* (*3*) state that with measured dose z and true dose x, a generally applicable approach to providing an unbiased assessment of the regression parameters is to replace the unknown x's in the disease model with their expectations $E(x \mid z)$. Thus in many problems, the primary goal of the statistician is to learn enough about the measurement model $f(z \mid x)$ and the true exposure distribution $f(x)$ that this expectation can be calculated from

$$E(x|z) = \frac{\int xf(z|x)f(x)\,dx}{\int f(z|x)f(x)\,dx}.$$

Here we assume that this has already been done, or even more generally, that we are able to characterize the full distribution $f(X \mid W)$, where W consists of the input data that determines the individual dose estimates (for example, W would be work history when doses are calculated by applying a job exposure matrix). The overall question that we discuss here is how to represent the uncertainty in the resulting ''Berkson'' dose distribution and how to incorporate this uncertainty (specifically shared uncertainty) into a dose–response analysis. If there are no shared errors (i.e., if the conditional distribution of X is independent between subjects), then the replacement of X with $E(x \mid z)$ in the dose–response analysis does give appropriate inference including both unbiased risk estimates (for models linear or close to linear in dose) and appropriate confidence intervals for the risk estimate. Our interest here, however, is in a problem in which there are shared errors in dose so that (for example) errors in job exposure affect dose estimates for everyone who worked at a particular job at a particular time.

Two problems to deal with then are how a dosimetry system should represent shared uncertainty and how the uncertainty should be included in the dose–response analysis. While confidence limits or standard deviations of each individual dose estimate may tell us something about the reliability of the dosimetry system, these by their nature cannot characterize the degree of sharing of dose error. One could attempt to produce a variance-covariance matrix for all the dose estimates, but this becomes extremely difficult if (as typical in occupational settings) the exposures are received over many different times. To date, dosimetry systems that really do attempt to characterize shared as well as unshared error [e.g. Hanford (*2*), Utah (*4*)] tend to do so by providing many replications from the dosimetry system rather than a single dose estimate. The variability of these replications of dose represents the cumulative impact of uncertainties in all dose-determining parameters, some of which have independent effects, but many of which affect many subjects' dose estimates simultaneously. For a simple example, if the source term in an environmental exposure (e.g. for down-winders) is known only up to certain level of certainty, then each of the replications of dose will use a different source term when the estimates are computed with that source term drawn from its likely ''prior'' distribution and combined with samples of all other uncertain parameters each drawn from their prior distributions. The variation in the random source term is therefore reflected in the variability of the dose estimates for all subjects simultaneously.

The conceptual underpinnings for our approach were further described in a recent paper by Stram and Kopecky (*2*). Denoting vectors with bold letters, we may view the true doses for the individuals in the study, X_i, as having been drawn from a distribution, $f(X_1, X_2, \ldots, X_N)$, of possible true doses that is consistent with the full set of input data, $\mathbf{W}$, known for the N subjects in the study. Here the term input data ($\mathbf{W}$) is used rather than measured dose (z above) because in complex dosimetry systems there may be many factors, $\mathbf{W}$, rather than a single best dose, z, that determine the possible doses that come out of a system for a given individual. For example, in systems as complex as that used in the Utah Thyroid disease study (*4*), the components of $\mathbf{W}$ could include place of birth and residence history through childhood, intake of milk, age during specific radiation releases, etc. of each study subject. The goal in building the dosimetry system is to provide a computer algorithm that samples ''possible doses'' from a distribution reflecting as closely as possible the conditional distribution of the true doses $f(X_1, \ldots, X_N \mid \mathbf{W})$ given the known input data.

Under this framework, the expected value, $\hat{X}_i$, of the true unknown dose may be estimated by its expectation over the samples from this distribution, which is $\hat{X}_i = E(X_i \mid \mathbf{W})$. The specific relationship between the true dose $\mathbf{X}$ and its estimate given the input data $\mathbf{W}$ can take different forms, and in our motivating application, this relationship is assumed to be multiplicative.

Use of Many Replications of Dose in Uncertainty Analysis

Assume that the dose–response relationship between the true doses $\mathbf{X}$ and the disease outcomes $\mathbf{D}$ is parameterized by a slope b of epidemiological interest and that there are one or more other parameters, $\mathbf{a}$, which need to be estimated in the likelihood. The complete distribution of all quantities of interest, namely the input data $\mathbf{W}$ for each subject, $\mathbf{X}$ and $\mathbf{D}$ for all subjects, is denoted $f(\mathbf{D}, \mathbf{X}, \mathbf{W} \mid \mathbf{a}, b)$. Since $\mathbf{X}$ is unknown, the likelihood is equal to the conditional distribution of

$$\begin{aligned} f(\mathbf{D}|\mathbf{W}, \mathbf{a}, b) &= f(\mathbf{D}, \mathbf{W}|\mathbf{a}, b)/f(\mathbf{W}) \\ &= \int f(\mathbf{D}, \mathbf{X}, \mathbf{W}|\mathbf{a}, b)\,d\mathbf{X}/f(\mathbf{W}), \end{aligned} \tag{1}$$

where $f(\mathbf{W})$ is the marginal distribution of input data in the cohort, assumed to not depend upon $\mathbf{a}$ or b. Here $\mathbf{a}$ and b are being treated as fixed but unknown parameters that determine the relationship between disease and dose but that do not affect the distribution of dose in the cohort. Note that $f(\mathbf{D} \mid \mathbf{W}, \mathbf{a}, b)$ is the likelihood function that we are truly interested in for statistical dose–response analysis: $f(\mathbf{D} \mid \mathbf{W}, \mathbf{a}, b)$ describes the distribution of disease, $\mathbf{D}$, given the observed values of all data that determine the dose estimates. This both involves the variability of true dose given estimated dose (through the integration) and reflects the variability in disease occurrence given true dose (incorporated explicitly into the integrand).

We assume that given $\mathbf{X}$ the distribution of $\mathbf{D}$ is independent of input data $\mathbf{W}$, i.e. that $\mathbf{W}$ functions only as a surrogate for $\mathbf{X}$ and does not itself have a direct effect on the risk of disease $\mathbf{D}$. Under this assumption (and the assumption that the distributions of X and W in the cohort do not depend on the dose–response parameters) we have

$$\begin{aligned} f(\mathbf{D}, \mathbf{X}, \mathbf{W}|\mathbf{a}, b) &= f(\mathbf{D}|\mathbf{X}, \mathbf{W}, \mathbf{a}, b)f(\mathbf{X}, \mathbf{W}) \\ &= f(\mathbf{D}|\mathbf{X}, \mathbf{a}, b)f(\mathbf{X}|\mathbf{W})f(\mathbf{W}), \end{aligned}$$

so that the observed likelihood (1) is

$$\begin{aligned} f(\mathbf{D}|\mathbf{W}, \mathbf{a}, b) &= \int f(\mathbf{D}|\mathbf{X}, \mathbf{a}, b)f(\mathbf{X}|\mathbf{W})\,d\mathbf{X} \\ &= E_{\mathbf{X}|\mathbf{W}}\{f(\mathbf{D}|\mathbf{X}, \mathbf{a}, b)\} \end{aligned} \tag{2}$$

where $E_{X|W}\{\cdot\}$ denotes taking the expectation (average) of the quantity

within the brackets { } conditional on the distribution of true dose given the observed data *W*. Maximum likelihood estimation of **a** and *b* then involves maximizing the average likelihood, $E_{X|W}\{f(\mathbf{D} \mid \mathbf{X}, \mathbf{a}, b)\}$, from a two-dimensional tabulation of estimates of this function. In the shared error framework, there is no independence between the subjects' estimates of dose, so that $f(X_1, X_2, \ldots, X_N \mid \mathbf{W})$ cannot easily be written as a product of terms $f(X_1 \mid W_1)f(X_2 \mid W_2) \cdots f(X_N \mid W_N)$. It follows that the entire multivariate distribution of $f(X_1, X_2, \ldots, X_N \mid \mathbf{W})$ is needed for Monte Carlo calculations.

Monte Carlo Simulations and Likelihood-Based Parameter Estimates and Confidence Intervals

Monte Carlo methods have been developed for integrating over high-dimensional probability distribution functions to develop inferences about model parameters or to make predictions. These methods have broad application in Bayesian statistics. Essentially, Monte Carlo simulations involve the repeated sampling from the specified distributions to form sample averages to approximate expectations (*5*). Here we use simulations to randomly select realizations from the distribution of true dose **X** given the measured data **W**.

Expression (2) allows us to calculate the value of the likelihood for fixed **a** and *b* by averaging the values of the likelihood $f(\mathbf{D} \mid \mathbf{X}, \mathbf{a}, b)$ over the random samples from the distribution of $f(\mathbf{X} \mid \mathbf{W})$.

The central estimate of *b* and the confidence intervals reflecting both the usual random error and the errors in dose can be derived from an integrated likelihood on a large number of simulations (e.g. 10,000) as follows: For each simulation, the likelihood of the model is determined by calculating the value of the likelihood at a large number of fixed values of *a* and *b* chosen to cover the maximum likelihood estimate and the range of possible confidence interval limits; the likelihoods at each of the points are then averaged over the simulations to approximate the integrated likelihood.

We note that model (2) is overly simplified since no additional covariates, **C** (such as age, sex or ethnicity), are assumed to modify the dose–response relationship between **D** and **X**. Introducing such covariates increases the dimensionality of the problem since additional parameters must be maximized out of the average likelihood. In fact, these covariates can also be part of **W** (if sex affects both dose distribution and background disease rate, for example) so long as the parameter linking **C** to **D** is distinct from the values of any parameter directly linking **C** and the dose distribution.

The maximum likelihood estimate (b_{MLE}) is then taken to be the value of *b* for which the profile likelihood, $L(b) = \max_{\mathbf{a}} f(\mathbf{D} \mid \mathbf{W}, \mathbf{a}, b)$, is at its maximum. Under appropriate regularity conditions the difference between $-2 \ln[L(b)]$ and $-2 \ln[L(b_{\mathrm{MLE}})]$ has an asymptotic (i.e. large sample) χ^2 distribution with one degree of freedom. The 90% confidence bounds were the values of *b* for which the χ^2 statistic is 2.7055.

For survival analysis using relative risk models, we can simplify the method further and average an alternative likelihood (often called the Cox partial likelihood), depending only on *b* rather than the full likelihood (depending on both **a** and *b*) over the distribution of dose. In this case the average (partial) likelihood, L(*b*), is found by averaging terms equal for a given realization of **X** to

$$\prod_i \frac{RR(X_i)}{\sum_{j \in \text{riskset}(i)} RR(X_j)},$$

where the product is over all the observed events (e.g. cancer deaths in our example) occurring in the cohort and the sum is over the risk set for event *i* (all subjects at risk at the age of the death for subject *i*). Here $RR(X_j)$ is the predicted relative risk for subject *i* given true dose **X**, which depends on X_i through the dose–response parameter *b*; see the example below for a specific model for RR_i.

Averaging the partial likelihood in this way is valid so long as the distribution, $f(\mathbf{X} \mid \mathbf{W})$, of the true dose given the determinants, **W**, of estimated dose is the same in all risk sets in the analysis. This homogeneity of **X** given **W** over all risk sets can safely be assumed to hold in situations where only a small proportion of the subjects in the study actually experience disease caused by their exposure (*6*). In addition, this also requires that censoring is non-informative about the dose distribution, implying, for example, that exposure is not causing (a significant number of) other deaths.

An Application

The methods described in this paper were developed for the analysis of a large multinational epidemiological study, the International Collaborative Study of Cancer Risk among Workers in the Nuclear Industry (the 15-Country Study) that was coordinated by the International Agency for Research on Cancer (IARC) (*7, 8*). In this study, annual records of external radiation doses were available from personal dosimeters for all workers. The Monte Carlo maximum likelihood method is illustrated here using data from one of the study facilities included in the 15-Country Study, the Oak Ridge National Laboratory (ORNL), since data from this cohort are publicly available on the U.S. DOE CEDR website (http://cedr.lbl.gov). Analyses of the risk of all cancers (excluding leukemia) related to external radiation dose were carried out. The ORNL cohort included in these analyses is described in detail elsewhere (*9*). Included were 5,345 workers, who were followed up between 1943 and 1984 for a total of 136,673 person-years. Following the protocol of the 15-Country Study, workers employed for less than 1 year and those who potentially received substantial doses from neutrons and internal contamination were excluded (*8*). Analyses included 225 deaths from all cancer excluding leukemia. Detailed results of ordinary analyses (i.e., analyses not taking account of uncertainties in dose estimates) of the ORNL cohort and other cohorts included in the 15-Country Study have been published (*7, 8*). An excess of all cancer mortality has been reported previously in analyses of the ORNL cohort (*8, 10, 11*).

A study of errors in recorded doses was conducted as part of the 15-Country Study to evaluate the comparability of recorded dose estimates across facilities and time and to identify and quantify sources of bias and uncertainties in dose estimates (*12*). The major sources of errors in external radiation doses were found to be dosimetry technology, exposure conditions in the workplace, and calibration practices. Systematic errors from these sources were characterized and quantified (*12*). For this, experiments were conducted to characterize the energy and geometry response of a sample of representative dosimeters (*13*), and the predominant conditions of exposure in different types of facilities were evaluated (*14*). Bias factors (B) were then developed using the error correction factors from the experiments, the predominant conditions of exposure, and energy- and geometry-specific conversion factors taken from the International Commission on Radiological Protection (*15, 16*). These bias factors were specific for each facility and for each period between times where there was a change in dosimeter use or other work practices that would affect the accuracy of the measurements. Estimates of the variances of the bias factors, which took into account uncertainties in the experimental results, the exposure conditions and the conversion factors, were also derived.

In our example we use as ''input data'' *W* the set of recorded doses, $W(t)$, for each year, *t*, that an individual worker was exposed. These recorded doses $W(t)$ are assumed to be related to the true doses $X(t)$ as $X(t) = W(t)/B$, where *B* is the bias factor for the specific type of dosimeter used in year *t*. Since the true bias factor *B* is uncertain, we sample *B* from a lognormal distribution with (log scale) variance corresponding to the uncertainties estimated by Thierry-Chef *et al.* (*12*) where the same bias factor applies to all individuals exposed at that given time and facility. Thus our dosimetry system specifies that the distribution of true dose is assumed to be lognormal around $W(t)$ with shared error term $1/B$.

The bias factors and uncertainties developed for the ORNL facility and used in the Monte Carlo simulations described below are presented in Table 1 for each period together with information on the distribution of annual doses for the cohort.

TABLE 1
Parameters of the Log-normal Distribution of Bias Factors Developed for ORNL and Summary of the Annual Recorded Dose Distributions by Period

	Parameters of the distribution		Summary statistics of annual dose (mSv)		
Year(s)	Geometric mean (β)	Uncertainty (κ)	Median	Mean	90th percentile
1943	2.048	1.601	0.50	1.33	3.51
1944–1952	1.353	1.698	1.30	3.04	7.50
1953–1979	1.201	2.464	0.96	1.78	3.94
1980–1997	1.137	1.708	0.75	1.03	2.00

Monte Carlo simulations were used in this example to randomly select a set of bias factors from the lognormal distribution for each facility and period. The resulting B's were then used to estimate annual ''corrected'' doses for each worker in the study by dividing their annual recorded dose by the appropriate sampled bias factor. Workers from the same facility were assigned the same correction factor for each year in the period and thus share a common Berkson error. Thus the correlation structure of the errors in the doses between individuals was preserved in the sampling process.

The sampling of the bias factors was repeated 10,000 times, which generated 10,000 data files containing corrected annual dose estimates for each individual in the cohort. The simulation size of 10,000 was selected because it resulted in stable estimates ($\leq$1% change) of the average b_{MLE} and confidence intervals over repeated samples.

Regression Model

Analyses of all cancer mortality (excluding leukemia) were conducted using the Cox proportional hazards model, stratifying on sex, age and socioeconomic status, to control for the potential confounding effects of these factors.

The parametric form of the model chosen for the evaluation was an additive relative risk model (*17*), also known as an excess relative risk (ERR) model in radiation epidemiology (*18*), which has the following basic form

$$\lambda[t \mid X(t)] = \lambda_0(t; \mathbf{a})[1 + bX(t)],$$

where $\lambda(t)$ is the hazard rate at time t, $\lambda_0(t; \mathbf{a})$ is the background hazard rate at time t, which may depend on one or more parameters $\mathbf{a}$, b is the slope parameter, and $X(t)$ is the individual's dose in sieverts (Sv), cumulated up to time t, allowing for a lag between exposure and cancer induction of 10 years. Thus RR_i in the definition of the partial likelihood above is equal to $[1 + bX_i(t)]$.

This model was applied to each of the ''simulated'' data files, and the partial likelihood was calculated at 100 values of the ERR parameter b. The values of the partial likelihoods at each value of b were then averaged over the simulations to obtain the average partial likelihood, from which the central estimate and confidence interval were derived.

For comparison, a conventional, unsimulated, analysis was also carried out by fitting an ERR model to the actual recorded dose estimates divided by the arithmetic mean of the distribution of the biases. This unsimulated model corrects for systematic bias but not for the uncertainties in the bias.

Implementation of the Procedure

A number of standard statistical packages exist for the fitting of Cox proportional hazards model to epidemiological data. However, these packages are highly computer intensive. For example, it took over 1 week to run 5000 simulations on this data set on a dedicated Unix workstation using EPICURE (*18*).

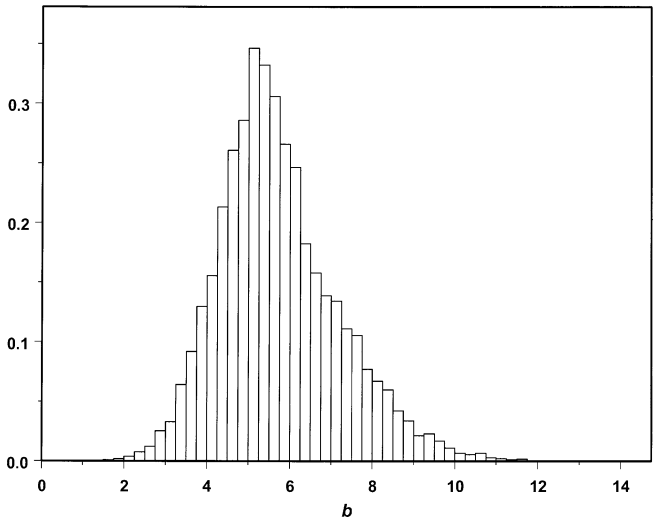

FIG. 1. Histogram of 10,000 estimates of b from Monte Carlo simulations.

To reduce the computational time needed to implement this approach on larger data sets, a computer program was written in FORTRAN to calculate the value of the model partial likelihood at each value of b, for each simulation. EPICURE was used to create a risk set for each of the cancer deaths in the cohort (i.e., the set of all subjects in the cohort that survived to at least the same age as the case) and to obtain the values of all analysis variables (including the stratification variables and each annual recorded dose) for each individual in each of the risk sets from the cohort data set. The risk sets and data for all subjects in the risk sets (including the deaths) were then exported from EPICURE. FORTRAN programs were then written and used to:

1. Conduct the Monte Carlo simulations, sampling the bias factors from the postulated distribution and creating 10,000 data sets, each corresponding to one realization of the simulated biases.
2. Calculate the partial likelihoods at each of the 100 preselected values of b for each of the 10,000 simulated data sets.
3. Average the partial likelihoods and find the b_{MLE} and confidence bounds.

Using these FORTRAN programs rather than EPICURE for the model fitting reduced the computing time from days to hours.

This research was exempt from institutional review for protection of human subjects because it used anonymized data.

RESULTS

A histogram of the b estimates from the 10,000 Monte Carlo simulations is presented in Fig. 1. The mean and median b of this distribution were similar (5.73 and 5.52 per Sv). The 5th and 95th percentiles of the distribution of the b estimates were 3.65 and 8.41, respectively. One might be tempted to report these values as the confidence interval accounting for the error in doses. However, the confidence interval derived in this way yields far too narrow bounds since they reflect only the shared uncertainty in dose and ignore the random error. They are simply descriptive statistics of the distribution of b.

Figure 2 illustrates the average partial likelihoods for each of the 100 chosen values of b. The average partial likelihood was maximized at b_{MLE} of 4.82 per Sv and the

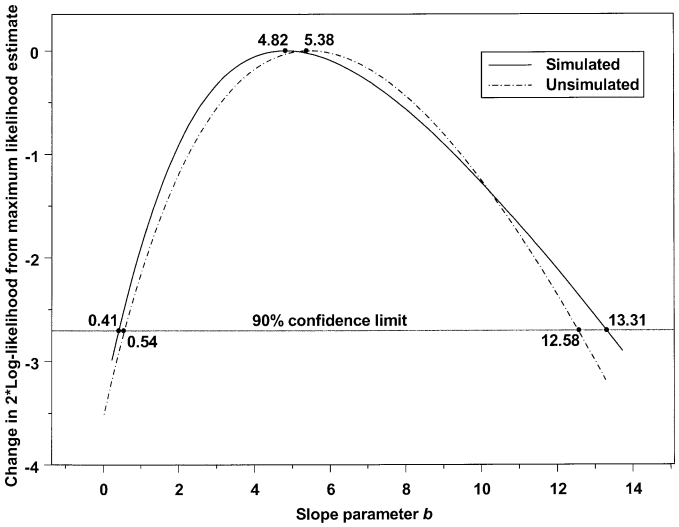

FIG. 2. Profile likelihood, maximum likelihood estimate and 90% CI for b from both the unsimulated data and the 1,000 simulations.

90% confidence interval ranged from 0.41 to 13.31. As would be expected, this confidence interval was wider than that derived from the conventional, unsimulated, analysis (90% CI = 0.54 to 12.58).

It is noted that the value of b ($b_{\mathrm{MLE}} = 4.82$) at which the average partial likelihood is maximized is lower than the b_{MLE} from the unsimulated analysis ($b_{\mathrm{MLE}} = 5.38$). However, we can see from the figure that the average partial likelihood is quite flat near its maximum and that both $b = 4.82$ and $b = 5.38$ give nearly the same log likelihood (i.e., we cannot distinguish between these two values).

DISCUSSION

The effect of shared errors in the estimates of exposures or doses is a common concern in studies of the role of occupational and environmental exposures in the etiology of chronic diseases such as cancer. This is because historical information on exposures or doses is often lacking in these studies and must be estimated for certain jobs or departments, particularly during the early period of the study. Shared errors may also occur in studies (e.g., studies of nuclear workers) where correction factors are used to adjust individual dose estimates for differences in measurement methods that have occurred over time and by study facility. The objective of this paper was to illustrate how one could apply the method of Monte Carlo maximum likelihood proposed by Stram and Kopecky (*2*) to derive maximum likelihood estimates of the slope and confidence intervals that reflect both ordinary statistical sampling error and errors arising from the exposure or dosimetry system.

We chose for this exercise to use a cohort study of nuclear workers at the Oak Ridge National Laboratory (ORNL). The primary source of errors in this dosimetry system was related to the response of historical dosimeters used to estimate radiation doses to the radiation fields encountered in the actual working environment of the study subjects. A bias correction factor was estimated to adjust for these responses, but there was uncertainty both in the exact magnitude of the response of the dosimeters and in the estimated exposure conditions encountered in the workplaces (*13, 14*).

It is interesting to note that, in this particular instance, the confidence intervals derived from this method were only slightly wider than those derived from the conventional analysis using the unsimulated data. This may not be the case, however, if the method is applied in other settings. In most occupational studies, the uncertainty in exposures is much larger than in studies of radiation workers, where extensive real-time individual monitoring of exposures has been conducted.

A few other papers have been published illustrating methods for handling uncertainty in exposure in the field of radiation epidemiology (*19–21*). The method developed by Gilbert (*19*) did not consider the shared errors as our method does. In our approach we are being somewhat more Bayesian about our dosimetry system than the more fully frequentist approach of Schafer *et al.* (*20*). Both methods treat the unknown true exposures as random variables but differ in their treatment of the uncertainty in the parameters that yield shared uncertainty. Here we are assuming that the dosimetry system that produces random realizations of dose directly incorporates the uncertainties in the bias factors through a Bayesian analysis in which these biases are treated as random variables coming from a prior distribution that has been determined in the study of errors in doses; in the approach of Schafer *et al.* (*20*), the data from the phantom studies would have been included simultaneously in the likelihood analysis to form a likelihood that combined both the observational data (cancer outcome) and the experimental data (observed and true doses). Essentially our approach separates the data used in the experimental assessment of the bias and uncertainty factors (B and K) from the data (survival times) that allow us to estimate the risk parameter, b. The full likelihood approach of Schafer *et al.* combines both data sets, so that some additional information about the bias factors would be (at least theoretically) gleaned from the risk analysis. For example, if one period and plant had higher than expected cancer deaths, then this observation would be used as information that would affect the bias correction factor for the corresponding dosimeter (raising it), something that is not being done in our method. However, it appears very likely that the amount of information regarding B in the survival analysis is far smaller than the information from the experimental exposures and expert estimation, so that the two methods should give very similar results concerning the distribution of bias factors (B) and by implication of b.

Lash and Fink (*21*) recently described methods for conducting sensitivity analyses to assess the impact of systematic errors of covariates used in epidemiological studies. Their approach, when applied to the problem of errors in

exposures here, recommends analyzing the cancer data using each replication of dose and then summarizing the resulting distribution of b, adding the variance of b from the simulation to the sampling variance of the estimate of b in the unsimulated analysis. This approach is related to the ''multiple imputation'' method of Rubin and coworkers (*22*). A complication to direct application of Rubin's multiple imputation method to the exposure measurement error problem is the need to condition, when simulating the exposures, not only on the data that are directly incorporated into the dosimetry system but also on the outcome (cancer) data in the analysis, so that a person with disease is given a different distribution of dose than a person without disease, all other factors being the same. In some cases, failing to condition on disease can result in large biases toward the null when each data set is analyzed separately. Interestingly, it is in the ''unshared'' error situation where this bias toward the null has its largest effect (*2*).

In the particular setting here, where sharing of errors is large and where the dose–response relationship is relatively weak, failing to condition on disease when drawing dose realizations is by itself not likely to cause serious difficulties in the multiple imputation method. However, another weakness of the multiple imputation method for these data has to do with the non-symmetrical shape of the log partial likelihood curve as a function of the dose–response parameter. The skewed shape of the log partial likelihood implies that non-symmetrical confidence intervals better capture the sampling variation in the data than do the symmetrical intervals given by either the Wald tests or the adjusted Wald tests generated using the multiple imputation method. Applying the Lash and Fink method to the data analyzed here modified the Wald-based 90% confidence intervals obtained using the conventional, unsimulated analysis from 0.54 to 12.54 to −1.22 to 12.10 after correction. This appears to overemphasize the uncertainty on the low side and underestimate the uncertainty on the high side, compared to the Monte Carlo likelihood-based method (90% CI: 0.41 to 13.31).

The b_{MLE} from our naïve unsimulated analysis ($b = 5.38$) was somewhat greater than the b_{MLE} from our Monte Carlo maximum likelihood analysis ($b = 4.82$). Some differences between the point estimate found by the Monte Carlo maximum likelihood approach compared to the conventional unsimulated analysis may be expected since the first of these approaches modifies the shape of the partial likelihood by averaging over a non-symmetrical (basically log-normal) distribution of possible values of true dose. This difference is expected to be very small, however. The difference noted here is most likely related to the distribution of the average partial likelihood function, which is quite flat at its maximum value; hence the ''apparent'' difference found in the ''best'' dose response using the two methods is not statistically meaningful.

In this application, despite the quite extensive ''sharing'' of errors in dose estimates at any one period, the net effect on the dose–response estimate and confidence limits was quite modest (the lower confidence interval decreased by about 24% and upper confidence limit increased by just 6%). This may indicate that no single period dominates the overall exposure to the workers studied here, so that the total amount of shared error in doses (summing over all periods) is smaller than it might otherwise have appeared to be. In other applications, the uncertainties of dose obtained during a single period or worker category may dominate total dose more completely than is the case here, and accounting for this would have a corresponding larger effect in the analysis.

We have described the Monte Carlo maximum likelihood method as a method for dealing with shared error. There are many simplifications available when only unshared error is present or when it dominates the error distribution. First, in the unshared case, the substitution method described by Thomas *et al.* (*3*) works very well for linear (or nearly linear) models (and can be extended to polynomial models). Computing the expectation

$$E(x\,|\,z) = \frac{\int x\ \Pr(z\,|\,x)\Pr(x)\ dx}{\int \Pr(z\,|\,x)\Pr(x)\ dx}$$

separately for each study subject (again z now refers to measured dose) is all that is required. The substitution method is a useful default approach to deal with the majority of measurement error problems (whether they are of classical, Berkson or other form) that do not involve shared errors so long as enough is known about the distribution of errors and true dose to make the integration feasible. As stated by Thomas *et al.* (*3*), the use of $E(x\,|\,z)$ may be regarded as turning the general measurement error problem into the Berkson problem, where both risk estimates and their estimated sampling uncertainties are unbiased (when the measurement errors are independent). This is not to say that unshared Berkson error does not cause problems. Random error in dosimetry always leads to loss of power to detect true dose response even when the estimates are unbiased, compared to having true dose available.

More generally the full likelihood

$$\prod_{i=1,N} \Pr(y_i\,|\,z_i)$$

can be computed by integration as

$$\prod_{i=1,N} \int \Pr(y_i\,|\,x_i)\ \Pr(z_i\,|\,x_i)\ \Pr(x_i)\ dx_i$$

far more readily than can the full likelihood when errors are shared (since the likelihood is then not the product of independent contributions that can be integrated separately). These simplifications mean that Monte Carlo maximum

likelihood methods rarely seem needed except to deal with the shared error problem.

ACKNOWLEDGMENT

Daniel Stram's work was supported by a grant from the U.S. Department of Energy, DE-FC01-07HS07010.

Disclaimer: The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the National Institute for Occupational Safety and Health.

Received: June 26, 2006; accepted: July 20, 2007

REFERENCES

1. B. Armstrong, The effects of measurement errors on relative risk regressions. *Am. J. Epidemiol.* **132**, 1176–1184 (1990).
2. D. O. Stram and K. J. Kopecky, Power and uncertainty analysis of epidemiological studies of radiation-related disease risk where dose estimates are based upon a complex dosimetry system; some observations. *Radiat. Res.* **160**, 408–417 (2003).
3. D. Thomas, D. Stram and J. Dwyer, Exposure measurement error: Influence on exposure–disease relationships and methods of correction. *Ann. Public Health* **14**, 69–93 (1993).
4. J. E. Till, S. L. Simon, R. Kerber, R. D. Lloyd, W. Stevens, D. C. Thomas, J. L. Lyon and S. Preston-Martin, The Utah Thyroid Cohort Study: analysis of the dosimetry results. *Health Phys.* **68**, 472–483 (1995).
5. C. Geyer, Estimation and optimization of functions. In *Markov Chain Monte Carlo in Practice* (W. R. Gilks, S. T. Richardson and D. J. Spiegelhalter, Eds.), pp. 241–258. Chapman and Hall, London, 1996.
6. R. L. Prentice, Covariate measurement errors and parameter estimation in a failure time regression model. *Biometrika* **69**, 331–342 (1982).
7. E. Cardis, M. Vrijheid, M. Blettner, E. Gilbert, M. Hakama, C. Hill, G. Howe, J. Kaldor, C. R. Muirhead and K. Veress, Risk of cancer after low doses of ionising radiation: retrospective cohort study in 15 countries. *Br. Med. J.* **331**, 77 (2005).
8. E. Cardis, M. Vrijheid, M. Blettner, E. Gilbert, M. Hakama, C. Hill, G. Howe, J. Kaldor, C. R. Muirhead and K. Veress, The 15-country collaborative study of cancer risk among radiation workers in the nuclear industry: Estimates of radiation-related cancer risks. *Radiat. Res.* **167**, 396–416 (2007).
9. M. Vrijheid, E. Cardis, M. Blettner, E. S. Gilbert, M. Hakama, C. Hill, G. Howe, J. Kaldor and K. Veress, The 15-country collaborative study of cancer risk among radiation workers in the nuclear industry: Design, epidemiological methods and descriptive results. *Radiat. Res.* **167**, 361–369 (2007).
10. D. B. Richardson and S. Wing, Greater sensitivity to ionizing radiation at older age: follow-up of workers at Oak Ridge National Laboratory through 1990. *Int. J. Epidemiol.* **28**, 428–436 (1999).
11. S. Wing, C. M. Shy, J. L. Wood, S. Wolf, D. L. Cragle and E. L. Frome, Mortality among workers at Oak Ridge National Laboratory. Evidence of radiation effects in follow-up through 1984. *J. Am. Med. Assoc.* **265**, 1397–1402 (1991).
12. I. Thierry-Chef, M. Marshall, J. J. Fix, F. Bermann, E. S. Gilbert, C. Hacker, B. Heinmiller, W. Murray, M. S. Pearce and E. Cardis, The 15-country collaborative study of cancer risk among radiation workers in the nuclear industry: Study of errors in dosimetry. *Radiat. Res.* **167**, 380–395 (2007).
13. I. Thierry-Chef, F. Pernicka, M. Marshall, E. Cardis and P. Andreo, Study of a selection of 10 historical types of dosemeter: variation of the response to Hp(10) with photon energy and geometry of exposure. *Radiat. Prot. Dosimetry* **102**, 101–113 (2002).
14. I. Thierry-Chef, E. Cardis, A. Ciampi, D. Delacroix, M. Marshall, E. Amoros and F. Bermann, A method to assess predominant energies of exposure in a nuclear research centre—Saclay (France). *Radiat. Prot. Dosimetry* **94**, 215–225 (2001).
15. ICRP, *Data for Use in Protection against External Radiation.* Publication 51, *Annals of the ICRP*, Vol. 17, No. 2/3, Pergamon, Oxford, 1987.
16. ICRP, *Conversion Coefficients for Use in Radiological Protection against External Radiation.* Publication 74, *Annals of the ICRP*, Vol. 26, No. 3, Pergamon, Oxford, 1997.
17. N. E. Breslow and N. E. Day, *Methods in Cancer Research*, Vol. 2. International Agency for Research on Cancer, Lyon, 1986.
18. D. L. Preston, J. H. Lubin, D. A. Pierce and M. McConney, *Epicure Users Guide.* Hirosoft, Seattle, 1993.
19. E. S. Gilbert, Accounting for errors in dose estimates used in studies of workers exposed to external radiation. *Health Phys.* **74**, 22–29 (1998).
20. D. W. Schafer, J. H. Lubin, E. Ron, M. Stovall and R. J. Carroll, Thyroid cancer following scalp irradiation: a reanalysis accounting for uncertainty in dosimetry. *Biometrics* **57**, 689–697 (2001).
21. T. L. Lash and A. K. Fink, Semi-automated sensitivity analysis to assess systematic errors in observational data. *Epidemiology* **14**, 451–458 (2003).
22. D. B. Rubin, *Multiple Imputation for Nonresponse in Surveys.* Wiley, New York, 1987.