Final Performance Report

Abstract

Recognition has been growing that radiation epidemiology studies need to assess the degree of measurement error in their radiation doses and a small number of attempts have been made to adjust radiation dose-response risk estimates for measurement error. However, the statistical assumptions and methods to correct for measurement error have had weaknesses that need to be addressed, and our work is a contribution to that end.

Radiation exposure measurement error from differential sources will be classified into two groups: unsystematic error and systematic error. Different models have be developed for these two types of measurement errors. A special type of systematic error caused by exposures below the minimum detection limit of dosimeters will be addressed separately from the other systematic errors. Occupational radiation exposure is often coded as zero when the exposure dose is below the minimum detection level. This leads to an underestimation of the doses received by individuals and can lead to overestimates of risk in occupational epidemiologic studies. The extent of the dose underestimation is increased with the magnitude of the minimum detection level (MDL) and the frequency of monitoring. We propose a Bayesian approach to estimate the actual dose and the dose distribution parameter when the observed dose is subject to censoring due to MDL. A Gibbs sampling algorithm is developed to implement the method. Simulation studies are used to evaluate the performance of the estimators.

We have developed a method to obtain an unbiased estimate of dose response relationship based on doses subject to minimum detection level. Our method has two steps: first we propose to a multiple imputation method to impute for missed doses and then to estimate the relative risk associated with exposure based on the average of the imputed missed doses. We develop a Gibbs Sampling method to impute for missed doses under two situations: (1) the radiation exposure is monitored annually and the observed annual dose is available; (2) the radiation exposure is monitored weekly but only the recorded annual dose obtained by the summation of all the observed weekly doses is available. The second situation arises in our example of the ORNL cohort data. It is more complicated because the number of weeks with BMDL doses is unknown. Then the average of the multiple imputed exposure realizations for each individual is used to obtain an unbiased estimate of the relative risk associated with exposure. Simulation studies are used to evaluate the performance of the estimators.

Given the absence of existing statistical methods to correct several type of measurement errors associated with occupational radiation exposure, this paper aims at assessing the impact of several type of measurement errors on dose response relationship so that an upper and lower bound of the true relative risk associated with exposure can be obtained. We consider random errors as well as doses subject to MDL. Simulation studies are used to examine the impact of these two types of measurement error on relative risk associated with radiation exposure. Several assumptions regarding the nature and magnitude of the errors in measuring true doses are made. The assumptions are hypothetical, but are chosen to approximate a range of possibilities likely to be encountered in actual occupational radiation studies. We apply the result from the simulation studies to a subset of the Oak Ridge National Laboratory (ORNL) (CEDR, 1999) data set in examining the range of true relative risk associate with radiation exposure.

Significant Findings

We propose Bayesian Monte-Carlo methods to adjust for the missed doses in two situations: (1) the radiation exposure is monitored annually and the observed annual dose is available; (2) the radiation exposure is monitored, say, weekly but only the recorded annual dose obtained by the summation of all the observed weekly doses is available. The second situation is more complicated because the number of weeks with below-MDL doses is unknown. We have shown that in the first case our method works well in estimating the dose distribution even when the censoring proportion is very high. In the second case our method works well in estimating individuals' doses when the censoring proportion is up to 60%. Furthermore, if we can obtain additional information for the dose distribution from an additional data set, then we are able to estimate individuals' dose well for up to 80% censoring. This suggests to us that when we design a study where the doses might be heavily censored, it is helpful to identify a small subgroup whose doses can be measured more precisely. Based on our method, the dose estimates can be viewed as the true dose measured with Berkson error. Therefore, our relative risk estimator is unbiased.

Usefulness of Findings

We apply our method to a subset of ORNL study population. The Oak Ridge National Laboratory (ORNL) is one of several facilities included in a large follow-up study of the health and mortality of workers at DOE facilities. We considered white males hired at ORNL between the opening of the facility in January, 1943 and the end of 1972 with follow up through 1984 and had worked for at least 3 years, which consist of 3,960 workers. We have found that the relative risks associated with radiation exposure were slightly overestimated

with the observed doses: the relative risk associated per 10 mSv increase in radiation doses changed from 1.018 to 1.017 when adjusting for missed doses due to doses below the minimum detection level. In another words, direct application of the observed doses gives fairly good estimates, perhaps because there is no dose effect with respect to all-cause mortality. A study having a strong exposure-mortality association would probably give a clearer assessment of the impact of below minimum detection level exposure correction.

It was noticed that in another major study of radiation exposure on health effects, the Hanford Study, the censoring due to below minimum may be severe in early years when the film badge was exchanged frequently. Our method can therefore be applied to this study. Our method is not limited to occupational radiation exposure problems but can be applied to a wide variety of environmental exposure-response data.

Scientific Report

1 Background

Estimates of external radiation dose, obtained from personal dosimeters, are used in epidemiologic studies of nuclear workers. A major objective of these studies is to provide a direct assessment of the carcinogenic risk of exposure to ionizing radiation at low doses and dose rates. In order to obtain an accurate and precise estimate of the risk of exposure, an accurate estimate of exposure is needed. However, the measurement of radiation exposure using personal dosimeters are subject to measurement errors. The sources of error identified by the National Research Council (NRC) Committee on File Badge Dosimetry in Atmospheric Tests are the following (Gilbert, Fix and Baumgartner, 1996). The first source of error is laboratory error including all errors introduced in film calibration, chemical processing, reading of optical densities, etc. A second source was identified as radiological error. Components of radiological error includes the energy spectrum (the failure of the dosimeter to respond accurately to all radiation energies to which personnel were exposed), wearing the dosimeter (the failure of a dosimeter worn on the torso to respond accurately to exposure coming from all directions), and backscatter associated errors (the overestimation that occurs when calibration is conducted in air rather than on phantom). The third source was environmental error including all error associated with the consequence of light, moisture, high temperature, etc., associated with the field environment. A fourth source of error was that resulting from converting recorded measurements to estimates of deep doses. Another source of error related to the fourth error is the error resulting from conversion of deep dose to organ doses. Modern dosimetry programs are usually designed to estimate "deep dose" rather than the

organ doses needed for epidemiologic purposes. Although it is possible to estimate factors for converting deep dose to organ doses (Gilbert, et al, 1996), these factors depends on the energy of the radiation and geometry. Among these four errors, the laboratory error and the environmental error are more likely due to random variation, can thus be modeled as random errors. The errors caused by the last two sources tend to be more systematic, thus may result in biased dose estimates.

Measurement error in dose estimates often biases estimated regression coefficients and may also result in underestimation of uncertainty and in distortion of the shape of the dose-response function. Methods for accounting for random error have been discussed by many investigators including Cochran (1968), Prentice (1982), Armstrong (1990), Clayton (1992) and Thomas et al. (1993). However, these methods require the magnitude (i.e., the distribution and the variation) of the measurement error to be known or estimable for validation studies. For occupational radiation exposure studies, in which this paper focuses on, these information is not available. Further, measurement error particularly random measurement error comes from uncertainty in single dosimeter readings. Occupation radiation exposure includes a series of exposure over time. Uncertainties in the estimates of the total exposure for individual workers, which are based on the sum of several dosimeter readings, vary by workers since workers can have different length of employment. The existing measurement error models mentioned above assume the magnitude of measurement error is the same across all the workers and therefore are not applicable here.

In addition to the sources of measurement errors described above, there is another special type of measurement error that is common in environmental studies, which is the occurrence of values below the minimum detection level (BMDL). The MDL is the lowest dose that a dosimeter can measure. The MDL of a dosimeter depends on how sensitive the dosimeter is and how experienced the dosimeter reader is.Uncertainty about the actual values below the minimum detection level (MDL) can bias or preclude subsequent statistical analysis. Consequently, the use of such data (type I left censored) for defining conditions and detecting trends or relationships can be compromised. A common practice in radiation studies is to record a zero for BMDL doses, which leads to an underestimate of the true dose. Conversely, if the MDL dose is recorded for BMDL doses, it will lead to an overestimate of the true dose. Either way will lead to a biased estimate of the dose-response relationship. Xue and Shore (2003, 2004) have developed methods to estimate the true dose and the true dose-response relationship when they are BMDL doses. However, their methods do not consider other sources of measurement error.

Given the absence of existing statistical methods to correct several type of measurement errors associated with occupational radiation exposure, this paper aims at assessing the impact of several type of measurement errors on dose response relationship so that an upper

and lower bound of the true relative risk associated with exposure can be obtained. We consider random errors as well as doses subject to MDL. Simulation studies are used to examine the impact of these two types of measurement error on relative risk associated with radiation exposure.

We apply our measurement error correcting method to the ORNL study. Checkoway et al. (1985) studied mortality of white males hired at ORNL between the opening of the facility in January, 1943 and the end of 1972 with follow up through 1977. The men all worked for at least thirty days, and there was no record indicating they had been employed at any other department at the facility. In this study, we extend the follow-up till 1984 and consider only workers who entered the laboratory after 1945 and had worked for at least 3 years, which consist of 3,960 workers. An individual's radiation dose of record at ORNL is based on pocket meters from 1943 to July 1944, film badges from then to 1975 and thermoluminescent dosimeters since 1975. The pocket meters were used in the early days of plant operations, particularly to obtain interim values between film badge readings, and continue to be used even today when high exposure potential is suspected (Parrish, 1982). The pocket meters were evaluated daily with a very low MDL (0.02 mSv), and the film badges were evaluated weekly from July 1944 to July 1956, when quarterly monitoring was initiated. The minimum detection level of the most sensitive film used at ORNL ranged from 0.10 to 0.30mSv (Kerr, 1994). A minimum detection level of 0.10 mSv was possible only if an experienced technician evaluated the exposed films with special care. During film badge exchange, when hundreds to thousands of films were read in large batches by technicians with widely varying level of experience, a minimum detection level of about 0.30 mSv is about as good as could be expected (Morgan, 1962). Annual monitoring was initiated in 1975 using thermoluminescent dosimeters (MDL=0.20 mSv). One complication of the ORNL dosimetry data is that only the annual doses are available in a computerized format. The annual dose before 1975 was obtained through the summation of the putative weekly or quarterly doses, which recorded zeroes for BMDL dose quantities. Most of the missing dose due to BMDL occurred from July 1944 till July 1956 when weekly measurements were made. With weekly measurements, the annual doses could be underestimated by as much as 15 mSv (0.30mSv/week X 50 weeks), although the chance that someone who has as high as 15mSv annual dose has all 50 weekly doses below MDL is very small. Both quarterly and annual monitoring have a much longer period to accumulate exposure compared to weekly monitoring, so there is a greater likelihood the accumulated dose will be above the MDL which makes the missed dose a less serious problem.

Several statistical methods were developed for adjusting the BMDL doses. Early works include Altshuler and Pasternack (1963) and Currie (1968) and more recent works include Newman et al. (1989), Taylor (1991), Mitchell et al. (1997) and ISO (2000). However, most of these methods are not applicable for the period 1944-1975 because the number

and identity of weeks or quarters with BMDL doses was generally unknown. Watkins et al. (1997) developed a method to compensate for likely zeroed BMDL doses in the ORNL annual recorded doses by a dose adjustment procedure, which involves a mathematical formula between observed doses and expected missed doses built from a small sample of daily pocket meter readings (where pocket meters have a high sensitivity to small doses). This approach ignored the sampling variation among doses and also depended heavily on the adequacy of the fitted model, which was built from the small sample of the data set that was computerized.

Not all the observed zero doses are BMDL doses. Some are legitimate zero doses since office workers tend to have no external exposure from radiation. Therefore, the observed zero doses are considered as the true zeroes if this person had more than 3 years of zero exposure or he had all zero doses and more than 75% workers of his department had zero exposure. For details, see Watkins et al. (1997). Except for legitimate zero doses and other special cases (for example, for workers whose film badges were not available, a plant median was used), the dose estimation method is applied to all the exposed years in order to obtain a cumulative dose of radiation for each subject. It is estimated that about 75% of the annual doses were censored due to BMDL and about 98% of the weekly doses were BMDL doses (Xue and Shore, 2003).

2 Specific Aims

2.1 Specific Aim 1: Distribution of the True Dose

Let Y_{ij} be the true dose for the ith subject at the jth year where $i = 1, \cdots, N$ and $j = 1, \cdots, n_i$. We assume Y_{ij} follows a gamma or a lognormal distribution since the exposure data tends to be right skewed.

Under the assumption of Gamma distribution, we assume the weekly dose follows $Gamma(\alpha, \beta)$ where α is the shape parameter and β is the scale parameter. We also assume the 50 weekly doses within a year (2 weeks of vacation are excluded) are independent and identically distributed. Therefore, the annual doses represented by Y should follow $Gamma(\alpha^*, \beta)$ where $\alpha^* = 50\alpha$.

Under the assumption of the lognormal distribution, we allow the repeated exposure for the same subject to be correlated. A random effect was therefore introduced to take into account the heterogeneity across subjects' exposures. Specifically, we assume that μ_i, the

mean of the logarithm of the true dose, $logY_{ij}$, is a random sample from a certain distribution. Then by conditioning on μ_i, the logarithm of the true dose Y_{ij} can be modeled as:

$$logY_{ij}|\mu_i \sim N(\mu_i, \sigma^2), \tag{1}$$

where we further assume $\mu_i \sim N(\mu, \tau^2)$. Doses from the same subject are thus correlated and the correlation is given by

$$Corr(logY_{ij}, logY_{ij'}) = \frac{\tau^2}{\sigma^2 + \tau^2}$$

for $j \neq j^{'}$.

The distribution of a weekly dose Y_{ijk} for $k = 1, \cdots, 50$ can then be numerically calculated, assuming (1). For simplicity, we approximate the distribution of Y_{ijk} by a lognormal distribution:

$$logY_{ijk}|\eta_i \sim N(\eta_i, r^2) \tag{2}$$

where η_i and r^2 are determined by matching the first two moments of Y_{ij} with $\sum_{k=1}^{K} Y_{ijk}$, so that

$$\eta_i = (log\frac{e^{2\mu_i}(e^{r^2} + \frac{K-1}{2})}{K^3} - r^2)/2 \tag{3}$$

and

$$r^2 = log(Ke^{\sigma^2} - \frac{K-1}{2}) \tag{4}$$

with $K = 50$. Similarly, the distribution of a quarterly dose can also be approximated by (2) with η_i and r^2 given by (3) and (4) and $K = 4$.

2.2 Specific Aim 3: Measurement Error due to MDL

We assume there was a minimum detection level and that doses below the MDL were coded as 0, as was true of the ORNL data. To estimate the true dose, we consider a simple situation first where the radiation exposure is only subject to uncertainty due to minimum detection level. A Gibbs sampling algorithm is developed to estimate the true dose. The Gibbs sampling approach is described in the following: under the Gamma distribution assumption,

1. Replace the censored annual doses (coded as 0) with samples from the bounded Gamma distribution, i.e.,$Y \sim Gamma(\alpha^*, \beta)|Y < MDL$;

2. Obtain the maximum likelihood estimator (MLE) of α^* and β after the censored doses are replaced;

3. Sample α^* and β based on the asymptotic multivariate normal distribution of the MLE of α^* and β obtained in 2.

With a noninformative prior on α^* and β, the posterior distribution is proportional to the likelihood function. The likelihood function can be approximated by a bivariate normal distribution through matching the first and the second derivatives at the mode. Therefore, the posterior distribution can be approximated by a bivariate normal distribution with the MLE as the mean and the inverse of the information matrix as the covariance matrix.

A simulation study has been used to evaluate the performance of the above method. In each simulation, we sampled an annual dose from $Gamma(\alpha^*, \beta)$ for each subject where $\alpha^* = 1.0$ and $\beta = 0.05$. These parameter values were set to approximate the shape of the dose distribution for our example, and also for simulation convenience. We set the annual observed doses to 0 if it is below the MDL. We let the MDL vary from 5,10,18,32 and to 46 so that the chances of missed doses were 20%, 40%, 60%, 80% and 90%, respectively. We checked to see how the method performs in estimating the actual annual doses under various censoring levels. The posterior sample for α and β and the actual annual dose for each subject with a zero observed dose (i.e., the simulated "actual" dose was below the MDL) was collected from the 5th iteration over 100 replicates. Then we obtained the mean of the parameters (α and β) and their 95% confidence intervals. For each subject with a zero observed dose, we used the mean of the posterior sample of annual doses for this subject as his estimated annual dose and defined the relative difference between the actual and the estimated doses for this subject as |estimated dose - actual dose|/actual dose$*$100% (For a subject whose annual dose is above the MDL, his observed dose is his actual annual dose and therefore his relative difference in annual doses is simply zero). We calculated the percentages of subjects that had a relative difference of below 10% and below 5% between their actual and estimated doses. We evaluated the performance of the method by first comparing how close the estimated distribution of annual dose was to the true dose distribution. This was done by checking how close the average of the sample means of α^* and β were to the true values and if the estimated 95% confidence interval retains its coverage probability. Second, we evaluated the proportion of persons for whom the relative differences between the estimated and the actual annual doses were below 10% and 5% and then average the proportions over 100 simulations. The simulation results are summarized in Table 1. Table 1 shows that the Monte Carlo method is able to estimate the distribution parameters well even with very high proportions of censoring, however, with a slightly underestimated variation. The proportion of relative differences below 10% (or 5%) hardly exceeds the percentage of whom had complete data (Table 1), indicating that this method cannot estimate individuals' doses well. This is what we expected because we do not have any information about the individuals' doses except that they are below the MDL. However, since the mean of the posterior sample for doses were used to estimate the actual doses, our estimators are unbiased estimators for

the actual doses while other methods such as "replacement of one half of the MDL" tend to either overestimate or underestimate the actual dose.

Under the lognormal distribution assumption, to impute the missed doses that are below the minimum detection level, a Gibbs Sampling method is also developed:

1. Replace the censored annual doses (coded as 0) with samples from the bounded lognormal distribution, i.e., $Y_{ij} \sim [lognormal(\mu_i, \sigma^2)|Y < MDL]$ based on current estimates of μ_i and σ^2;
2. Update μ_i by sampling from its posterior distribution, conditioning on the complete doses and current estimates of μ, σ^2 and τ^2;
3. Update μ by sampling from its posterior distribution, conditioning on the current estimates of $\mu_i, i = 1, \cdots, N$ and τ^2;
4. Update σ^2 by sampling from its posterior distribution, conditioning on the complete doses and current estimates of $\mu_i, i = 1, \cdots, N$;
5. Update τ^2 by sampling from its posterior distribution, conditioning on the current estimates of $\mu_i, i = 1, \cdots, N$ and μ.

Repeat steps 1-5 until convergence. Then we obtain an imputed dataset of complete doses. We repeat the process M times and the average of the M imputed doses is used to estimate the true dose.

Doses in the early years were evaluated on a weekly basis at ORNL. The doses below the MDL were coded as 0. When weekly doses are available, we know the weeks with BMDL doses. We can then estimate the censored weekly doses using the method described above. However, in the early years the number and identity of the weeks with censored doses are unknown. Only the annual doses obtained by the summation of the censored weekly doses are computerized. Every subject may have had BMDL doses for some weeks, which were recorded as zero.

Let Y_{ijk} denote the true weekly dose for individual i at jth year and kth week and $Y_{o_{ijk}}$ denote the corresponding observed dose. Then $Y_{o_{ijk}} = Y_{ijk}$ if $Y_{o_{ijk}} \geq$ MDL and $Y_{o_{ijk}} = 0$ otherwise. In early years, only the annual sum of $Y_{o_{ijk}}$'s is available, that is to say, only $Y_{o_{ij}} = \sum_{k=1}^{50} Y_{o_{ijk}}$ is available. Let $N_{u_{ij}}$ represent the number of weeks that the weekly doses are equal to or above the MDL for individual i at jth year, $N_{l_{ij}} = 50 - N_{u_{ij}}$ is the number of weeks that the doses are below the MDL. The prior distribution of N_u is binomial with

size 50 and success rate $P(Y_{ijk} > MDL)$. A Gibbs Sampling method to estimate the missed doses is developed as follows:

1. Sample $N_{u_{ij}}$ from $[N_{u_{ij}}|\text{observed doses}, \mu_i, \sigma^2]$, i.e., its posterior distribution, conditional on the observed doses and the current estimates of μ_i and σ^2;
2. Sample $N_{l_{ij}}$'s weekly doses from the bounded lognormal distribution;
3. Adjust the observed dose by adding the sampled weekly doses from Step 2. to the observed annual dose;
4. Update μ by sampling from its posterior distribution, conditioning on the current estimates of $\mu_i, i = 1, \cdots, N$ and τ^2;
5. Update σ^2 by sampling from its posterior distribution, conditioning on the complete doses and current estimates of $\mu_i, i = 1, \cdots, N$;
6. Update τ^2 by sampling from its posterior distribution, conditioning on the current estimates of $\mu_i, i = 1, \cdots, N$ and μ.

Repeat steps 1-6 until convergence. Then an imputed data set of complete doses is obtained. The process is repeated M times and the average of the M imputed doses is taken as the estimate of the true dose.

A simulation study is also used to evaluate the performance of the method. Since it is similar to the result presented in Table 1, it is omitted here.

2.3 Specific Aim 2: Measurement Error Model in Dose Estimates

The observed dose Z_{ij} for the ith subject at the jth year is subject to a random measurement error. Assuming a classical measurement error model on the log scale, i.e., a multiplicative random measurement error,

$$logZ_{ij} = logY_{ij} + \epsilon_{ij} \tag{5}$$

where ϵ_{ij}'s are independent random measurement error which are assumed to follow $N(0, \sigma_\epsilon^2)$. Under the measurement error model of (5), each weekly dose Y_{ijk} for $k = 1, \cdots, 50$ is subject to a multiplicative measurement error $e^{\epsilon_{ij}}$, i.e., the observed weekly dose, denoted by Z_{ijk} is $Z_{ijk} = Y_{ijk}e^{\epsilon_{ij}}$.

2.4 Specific Aim 4: Dose-response relationship

The primary interest is to estimate the relationship between survival and the true radiation exposure history. Let $Y(t)$ denote the true dose at time t, and let $\overline{Y(t)}$ denote the history up to time t, $\{Y(u), u \leq t\}$. We use a time dependent Cox proportional hazards model (Cox, 1972) to model this relationship, i.e.,

$$\lambda(t|\overline{Y(t)}, \overline{X(t)}) = \lambda_0(t) e^{\theta_1 \sum_{u \leq t-g} Y(u) + \theta_2 X(t)} \tag{6}$$

where $\overline{X(t)}$ is the history of other covariates $X(u)$ for $u \leq t$, which are assumed to be observed without error. Generally there is a lag time, say, g, before a disease caused by the radiation exposure becomes manifest.

2.5 Specific Aim 5: Estimation of dose-response relationship for radiation exposure measured with measurement error

As discussed earlier, when the radiation exposure $Y(t)$ is measured with a certain minimum detection level, simply using the observed doses leads to a biased estimator of θ_1. To summarize our estimation method, first we obtain samples from the estimated distribution of the true dose conditioning on the observed doses; let y_r^* denote a sample of the estimated distribution of $Y|Y_o$, then we use $z = \sum_{r=1}^{M} y_r^*/M$ to estimate the true dose. Provided M is large enough,

$$Z \approx E(Y|Y_o)$$

and thus Z can be viewed as a dose measured with a Berkson error so that

$$Y = Z + \epsilon$$

where ϵ has mean 0 and a constant variance. Therefore, the estimated cumulative exposure based on our estimation method, $\sum_{u \leq t-g} z(u)$ is essentially a dose with a Berkson error. We further assume ϵ follows a normal distribution. Then following the argument of Prentice (1982) and Pepe et al. (1989), the estimate of θ_1 based on our dose estimate is unbiased.

A simulation study is used to evaluate the performance of the relative risk estimator. In each simulation, we first sample the mean parameter μ_i from a normal distribution with mean $\mu = 3.5$ and variance $\tau^2 = 0.6$ for $i = 1, \cdots, N$. Then given each μ_i, we sample y_{ij} for $j = 1, \cdots, n_i$ from a lognormal distribution such that $logy_{ij} \sim Normal(\mu_i, 0.4)$. Therefore, we have $EY_{ij} = 54.5$ millirems= 0.545 mSv, $VarY_{ij} = 0.8013$ mSv^2 and $Corr(logY_{ij}, logY_{ij'}) = 0.6$. For simplicity, we assume there are no other covariates except the dose. Then based on

the true dose level and an assumption on the lag g, we generate survival times using model (6) with the relative risk parameter $\theta_1 = 0.02$ per 10 mSv increase in dose; the baseline hazard is set to be a constant, 0.01333 so that the probability of surviving over 10 years without being exposed to radiation is about 87.5%. The parameters are chosen so that the distribution of the simulated data is close to the data in the example.

With a chosen MDL, we set the annual observed doses to be 0 if it is BMDL. We let the MDL vary from 0.10mSv, 0.18mSv, 0.30mSv, 0.50mSv, and to 0.80mSv so that the chances of missed doses were 10%, 25%, 45%, 65% and 80% respectively. We checked to see how the method provided in Section 2.1 performs in estimating the true relative risk under various censoring levels. For each observed zero dose, we took the sampled doses at the 5th iteration over 10 replicates and used the average as the estimate of the missing doses. Then we estimated the relative risk associated with the radiation exposure based on the "complete dose". We repeated the simulation 300 times; Table 2 summarizes the results.

Table 2 shows that when the observed doses, which are left censored (because they are BMDL), are directly applied, the relative risk associated with the exposure (i.e., θ_1 in model (6)) is generally underestimated. However, the underestimation becomes relatively severe only when the proportion of BMDL annual doses is 30% or higher. The relative risk estimates based on the imputed doses perform much better and generally give an unbiased estimator of the true relative risk. When the proportion of censoring goes up to 80%, the model with the imputed doses slightly underestimates the true relative risk by 3%, but the observed dose statistic underestimates by 19%. The estimated standard error for the relative risk estimates based on the observed doses also is generally too small; the model with imputed dose gives a better estimator of the standard error except when the censoring is quite high (80% or higher).

For the subset of ORNL data, a total of 10 imputed datasets of doses were generated, adjusted for missing doses due to BMDL. Then the average of the imputed doses for each worker was taken as the estimated dose in order to calculate the relative risk associated with the cumulative radiation exposure. We evaluated the association of radiation exposure with all-cause mortality using a time-dependent Cox proportional hazards model. The Cox model controlled for sociodemographic factors as well, although data on smoking, chemical exposures, medical exposures to ionizing radiation, and cancer mortality were not available. We used the sociodemographic variables such as age, birth year, pay code (monthly/nonmonthly) and active/inactive worker status. Pay code was used as an indicator of socioeconomic status with monthly paid workers to be on a higher socioeconomic level. Active worker status was considered because workers who continued employment, and consequently exposure, tend to be healthier. A lag of 10 years was used.

Table 3 presents the estimates of the relative risk associated with exposure and other covariates when the observed doses and the imputed doses were used. Table 3 shows that the relative risks associated with radiation exposure were slightly overestimated with the observed doses: the relative risk associated per 10 mSv increase in radiation doses changed from 1.018 to 1.017 when the imputed doses were used. Similarly, very close results were obtained for other covariates as well, indicating that in this study, direct application of the observed doses gives fairly good estimates, perhaps because there is no dose effect with respect to all-cause mortality. A study having a strong exposure-mortality association would probably give a clearer assessment of the impact of BMDL exposure correction.

When the radiation exposure is both subject to random error and minimum detection level, it is complicated to use an analytical approach to obtain an unbiased estimator of the relative risk association with radiation exposure. A simulation study is then proposed to evaluate the magnitude of bias and loss of efficiency. The development of this simulation study is undergoing and is expected to be finished by the end of May.

Table 1. The estimation of dose distribution parameters and individuals's true dose when the exposure is monitored annually with a MDL: results based on 100 simulations where the annual dose is assumed to be Gamma distributed with $\alpha^* = 1.0$ and $\beta = .05$.

% of below[1] MDL doses	α^*		β		Relative diff. in annual dose[3]	
	average of sample mean[2]	coverage prob.(%) of the 95% CI	average of sample mean	coverage prob.(%) of the 95% CI	average of % < 10%	average of % < 5%
20%	1.128	98.0	.056	92.0	80.1	80.0
40%	1.143	99.0	.057	90.0	64.5	62.6
60%	1.138	100.0	.053	94.0	45.6	42.9
80%	1.102	100.0	.043	98.0	26.5	23.2
90%	1.191	100.0	.056	89.0	17.3	13.7

Note: 1. The MDL is chosen such that the proportion of doses below that level is from 20% to 90%. 2. In each simulation, we obtain a sample of 100 points of (α^*, β). The sample mean is then taken as the estimate of the parameters. We also constructed the 95% confidence interval based on the normal assumption of the sample mean. 3. For each subject with a below MDL dose, we obtain a sample of 100 points for his doses. We then use the mean of the sample to estimate his actual annual dose. The relative difference in annual doses is calculated as |estimated dose-actual dose|/actual dose$*$100%. For subjects whose annual doses are above the MDL, their relative differences in annual doses are zero. Then the average across subjects of the relative difference is calculated.

Table 2. The estimation of the relative risk based on the observed dose and the adjusted dose when the exposure is monitored annually and a fraction of exposures are BMDL: results based on 300 simulations ($\theta_1 = 0.02$)

% of below	Based on Observed doses		Based on Imputed doses		
MDL doses	average[1] of $\hat{\theta_1}$	average of $\hat{se}(\hat{\theta_1})$	average of $\hat{\theta_1}$	average of $\hat{se}(\hat{\theta_1})$	Sample se
10%	.0196	.00131	.0200	.00133	.00133
30%	.0188	.00125	.0202	.00134	.00138
45%	.0174	.00117	.0199	.00133	.00135
65%	.0166	.00113	.0198	.00136	.00142
80%	.0162	.00117	.0193	.00143	.00238

Note: 1. The average is taken over 300 simulations.

Table 3. Comparison of parameter estimates in the Cox model (6) for the association between radiation exposure and all cause mortality using observed radiation doses and imputed radiation doses for a subsample of ORNL study

Variables	observed doses			imputed doses		
	$\hat{\theta}_1$	$\hat{se}(\hat{\theta}_1)$	p-value	$\hat{\theta}_1$	$\hat{se}(\hat{\theta}_1)$	p-value
radiation dose (in 10 mSv)	.018	.011	.088	.017	.010	.097
current employment	-.561	.107	<.001	-.561	.107	<.001
age in years	.089	.004	<.001	.089	.004	<.001
paycode (monthly)	-.610	.095	<.001	-.609	.095	<.001

Publications

Xue, X. and Shore, R.E.(2003) Methods for estimating occupational radiation doses subject to minimum detection levels. *Health Physics*, **84(1)**: 61-71.

Xue, X., Shore, R.E., Ye, X. and Kim, M.Y. (2004) Estimating the dose response relationship for occupational radiation exposure measured with minimum detection level. *Health Physics* In press.

Xue, X., Kim, M.Y. and Shore, R.E. Estimating risk associated with occupational radiation exposure when radiation exposure were measured with measurement error and minimum detection level. In preparation.

References

Please see the attached manuscripts.

Equipment Inventory

A Dell Inspiron 7500 Notebook was bought in March of 2000; Dr. Xue took it with her when she moved to Albert Einsten College of Medicine in May of 2003;

A Dell Dimension 4100 Computer was bought in February of 2001; Dr. Xue left it at NYU School of Medicine when she moved to Einstein.

Final Invention Statement

No inventions were conceived under the grant.