Paper

ESTIMATION OF HEALTH RISKS ASSOCIATED WITH OCCUPATIONAL RADIATION EXPOSURE: ADDRESSING MEASUREMENT ERROR AND MINIMUM DETECTABLE EXPOSURE LEVEL

Xiaonan Xue,* Mimi Y. Kim,* and Roy E. Shore†

Abstract—**Occupational exposures are subject to several types of measurement errors. This paper considers two of the most common types of measurement error associated with occupational exposures: the error due to below minimum detection level and doses due to random measurement error. Doses are often recorded as zero when the exposure is below the minimum detection level. Values that are below the minimum detection level and are entered as zero lead to underestimation of the true exposure and can result in either an overestimate or underestimate of risk associated with the exposure. Random measurement error leads to an inefficient and attenuated estimate of risk associated with exposure. However, the levels of bias and inefficiency that can result from the simultaneous presence of both types of measurement error have not previously been studied. In addition, the impact of these measurement errors on the type I error and type II error for an exposure-response effect is unclear. Since the magnitude of the random error associated with cumulative exposure may vary with individuals and across time within an individual, traditional methods to correct for random measurement errors are not applicable here. Further, correcting errors for minimum detectable levels and random errors simultaneously is too complex for analytical solutions. Therefore, this paper uses simulation studies to quantitatively evaluate the magnitude of the bias, inefficiency, and type I and type II errors associated with them. The simulation results are applied to a sample of historical occupational radiation exposure data from the Oak Ridge National Laboratory. We conclude that after taking into consideration random measurement error and missed doses due to falling below the minimum detection level, radiation exposure is not significantly associated with all-cause mortality in that cohort.**

Health Phys. 91(6):582–591; 2006

Key words: analysis, risk; analysis, statistical; biostatistics; dosimetry, errors

INTRODUCTION

ESTIMATES OF external radiation dose obtained from personal dosimeters have been used in numerous epidemiologic studies of nuclear workers (e.g., Cardis et al. 2005; Dupree-Ellis et al. 2000; Gilliland et al. 2000; Ritz 1999; Gilbert et al. 1993; Wiggs et al. 1994). 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. Unbiased, precise exposure estimates are needed for an adequate assessment of risk; however, measurement of radiation exposure using personal dosimeters is subject to measurement errors. The sources of error identified by the National Research Council (NRC) Committee on Film Badge Dosimetry in Atmospheric Tests are the following (Gilbert et al. 1996). The first is laboratory error, including errors introduced in film calibration, chemical processing, reading of optical densities, etc. A second source is radiological error, including the energy spectrum (the failure of the dosimeter to respond accurately to all radiation energies to which personnel are exposed), dosimeter placement (e.g., the failure of a dosimeter worn on the torso to respond accurately to exposure coming from all directions), and backscatter associated errors. The third source is environmental error including errors associated with the consequence of light, moisture, high temperature, etc., in the field environment. A fourth source of error is in trying to estimate appropriate organ/tissue dose from dosimetric readings that are usually designed to estimate "deep dose," typically, the energy absorbed at a depth of one cm from the body surface. Although it is possible to estimate factors for converting deep dose to organ doses (Gilbert et al. 1996), these factors depend on the energy and geometry of the radiation source. Among these four errors, the laboratory error and the environmental error are more likely due to random variation and can thus be modeled as random errors. The errors caused by the other two sources tend to be more systematic and thus may

* Division of Biostatistics, Department of Epidemiology & Population Health, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461; † Radiation Effects Research Foundation, Hiroshima, Japan.

For correspondence contact: Xiaonan Xue, Division of Biostatistics, Department of Epidemiology & Population Health, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, or email at xxue@aecom.yu.edu.

(*Manuscript accepted* 11 *April* 2006)
0017-9078/06/0

result in biased dose estimates. Another special type of systematic error that is common in environmental studies is the occurrence of values below the minimum detection level (MDL). The MDL is the lowest dose that a dosimeter can measure, because of background "noise" and limited sensitivity. The MDL of dosimeters prior to modern automated readers depended on both the sensitivity of the dosimeter and the experience of the dosimeter reader. Uncertainty about the actual values below the minimum detection level (BMDL) can bias subsequent statistical analysis. Consequently, the use of such data, which is equivalent to type I left censored data, for defining conditions and detecting trends or relationships can be problematic. 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 approach will lead to a biased estimate of the dose-response relationship. The differences between systematic errors and random errors are that (1) random error is independent from the true dose and therefore leads to an unbiased, albeit imprecise, estimate of the true dose, and systematic error may be dependent on the true dose and lead to a biased estimate of the true dose; (2) random error is different for all workers but systematic error may be the same for a specified group of workers. This paper focuses on random errors and errors due to BMDL. Other types of systematic error tend to be study-specific and a thorough review of the particular dosimetry practices would be needed to estimate the magnitude of those systematic biases (Gilbert et al. 1996; Gilbert 1998). Though we do not consider other types of systematic errors here, our simulation method can easily be generalized to include them.

Random dose measurement error often biases estimated regression coefficients, usually in the direction of attenuating the dose-response slope. When it is not taken into account, it also results in underestimation of uncertainty and may distort the shape of the dose-response function (Gilbert 1998). 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 either known or estimable through validation studies. In occupational radiation exposure studies, validation studies are generally not available and the magnitude of the random measurement error is often unknown, although the range of the magnitude of random measurement error might be obtained via calibration data (Gilbert and Fix 1996). Further, occupational radiation exposure includes a series of exposures over time. Random measurement error comes from uncertainty in every dosimeter reading. Therefore, uncertainties in estimates of the cumulative exposure, which are based on the sum of single dosimeter readings (such as annual dosimeter readings) over the length of exposure, may vary by worker since workers have different lengths of employment and may vary by time since exposure. The existing measurement error models mentioned above all assume the distribution of the random measurement error to be the same for all the workers and do not change with time and therefore are not applicable here.

Xue and Shore (2003) and Xue et al. (2004) have developed methods to estimate the true dose and the dose-response relationship when there are BMDL doses. However, their methods do not consider random measurement errors. Since the random measurement errors may mask the dose-response relationship and thus have a critical impact on the risk estimates associated with occupational radiation exposures, it is essential to take measurement errors into consideration. Currently, there are no existing statistical methods available to evaluate the impact of these two types of measurement error simultaneously and to correct for the errors for the purpose of evaluating the risk associated with occupational exposures. Mitchell et al. (1997) developed a Bayesian approach to estimate true dose for occupational radiation exposure when it is subject to random measurement error and error due to MDL; however, the exposure-response relationship was not estimated. Richardson and Ciampi (2003) studied the effect of random measurement error and error due to minimum detection level on a continuous or a binary response, but the exposure they studied is a one time exposure rather than cumulative exposures. To estimate risk associated with cumulative exposures, a time-dependent Cox proportional hazards model is generally used. Complicated statistical computations are required in the dose-response modeling to take into account both random error and error due to BMDL, and to allow the magnitude of the uncertainty associated with cumulative exposure to vary across individual workers and vary by time since employment within an individual. Thus, the application of such a method, if developed, will be difficult. Given the absence of easily implemented analytical statistical methods to correct for the several types of measurement errors associated with cumulative radiation exposure, this paper uses simulation studies to assess the impact of the two types of measurement error on the dose-response relationship. We then apply the simulation method to a subset of the Oak Ridge National Laboratory (ORNL) (CEDR 1999) data set to reexamine the true relative risk of all-cause death associated with radiation exposure

after "correcting" for random errors and errors due to BMDL. Below we will first present the risk model and measurement error models. Next we conduct a simulation study for hypothetical data sets, and finally we use simulation studies to re-evaluate the risk model specifically for the ORNL data set.

MODEL

A log linear Cox proportional hazards model is used with the cumulative exposure $Y(t)$ as a time-dependent variable and other possible confounding variables (Xue et al. 2004). The primary interest is to estimate the relative risk for mortality associated with cumulative radiation exposure, e^{β}. As discussed earlier, when the radiation exposure $Y(t)$ is measured with random measurement error and a certain minimum detection level, simply using the observed doses leads to a biased and inefficient estimator of β (i.e., the variance of the estimate of β is larger when the doses are measured with random measurement error). The amount of bias and loss of efficiency depends on the true dose distribution, the magnitude and the distribution of the random measurement error, and the minimum detection level. In the following, we describe our models of the true dose distribution, random measurement error, and our assumption for the minimum detection level.

The dose distribution proposed in Xue et al. (2004) is used here, i.e., 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 lognormal distribution with the mean parameter μ_i. Thus, each subject has its own mean level of exposure, and we further assume $\mu_i \sim N(\mu, \tau^2)$. Doses from the same subject are therefore allowed to be positively correlated; i.e., if a subject received a high dose this year, he is more likely to receive a high dose next year at the same job.

A lognormal distribution is chosen to model the true dose because of the skewed nature of the dose distribution and the convenience of introducing a multiplicative lognormal random error. It should be noted that the lognormal distribution is only used to model positive exposures. In occupational studies, there are jobs associated with no radiation exposure. The legitimate zero doses are usually assessed based on job categories and other criteria (Watkins et al. 1997). The observed dose Z_{ij} for the ith subject at the jth year is subject to random measurement error. Assuming a classical measurement error model on the log scale, i.e., a multiplicative random measurement error,

$$logZ_{ij} = logY_{ij} + \epsilon_{ij}, \quad (1)$$

where ϵ_{ij} are independent random measurement errors following a normal distribution with mean 0 and variance σ^2. Our random measurement error model is similar to the third random measurement error model proposed by Gilbert and Fix (1996) and Gilbert (1998). In their model, the estimated dose Z from a single reading (could be a weekly reading or an annual reading) has the true dose Y as the mean and variance K_3Y^2; while in our model, the estimated dose Z from a single reading has mean $Ye^{\sigma^2/2}$ and variance $Y^2(e^{2\sigma^2} - e^{\sigma^2})$. When σ^2 is small, $e^{\sigma^2/2}$ is close to one. Therefore, our estimator is close to being an unbiased estimator of the true dose when the error is small. The extension of our model is that we allow the repeated exposure measurements to be correlated with each other while Gilbert assumed they are all independent. It should also be noted that the random measurement error model used here is a classical measurement error model. Contrary to the Berksons measurement error, the classical measurement error tends to attenuate the relative risk estimates. In addition to the random measurement error, the dosimeter readings have to be above the minimum detection level. In other words, only Z^+ is observed where

$$Z^+ = \begin{cases} Z & \text{if } Z \geq \text{MDL} \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

In the early years at ORNL, doses were evaluated on a weekly basis, so we simulated this scenario. Assume a worker generally works 50 wk each year, then the distribution of a weekly dose Y_{ijl} for $l = 1, \cdots, 50$ (for quarterly, $l = 1, \cdots, 4$) can be approximated by a lognormal distribution with the mean parameter and the variance parameter given in Xue et al. (2004). Assume the following random measurement error model:

$$logZ_{ijl} = logY_{ijl} + \epsilon_{ijl}, \quad (3)$$

where ϵ_{ijl} are independent random measurement errors following the same distribution as ϵ_{ij}. Then each weekly dose Y_{ijl} is subject to a multiplicative measurement error $e^{\epsilon ijl}$.

Due to the MDL, only Z^+_{ijl} is observed where $Z^+_{ijl} = Z_{ijl}$ if $Z_{ijl} \geq$ MDL and $Z^+_{ijl} = 0$ otherwise. Only the annual doses obtained by the summation of the observed weekly doses were computerized, i.e., only $Z^+_{ij} = \Sigma^L_{l=1} Z^+_{ijl}$ was available at ORNL in early years.

Direct application of the observed dose Z^+ in the Cox model leads to biased estimates of β as well as the variance associated with the estimator. To obtain an accurate and efficient estimator of β analytically, for example, using the maximum likelihood approach, one has to integrate the likelihood function over the measurement error distribution as well as the distribution of the

missed doses due to MDL. The measurement error associated with the cumulative exposure $\Sigma^{t}_{i=1} Y_{ij}$ at a given year t since entry is $\Pi^{t}_{i=1} e^{\epsilon ij}$. The distribution of $\Pi^{t}_{i=1} e^{\epsilon ij}$ is lognormal with mean parameter 0 and variance parameter $t\sigma^2$. Thus, the distribution of the measurement error associated with the cumulative exposure changes with t. Assuming doses are monitored annually, the missed dose due to MDL for the cumulative exposure $\Sigma^{t}_{i=1} Z_{ij}$ is $\Sigma^{t}_{i=1} Z_{ij} I(Z_{ij} < \text{MDL})$, whose distribution also varies with time t. Therefore, the implementation of an analytical method to obtain an accurate and efficient estimator of β is almost intractable, so simulation studies were used to quantitatively assess the impact of the errors on statistical inference for dose-response relationships under various levels of these two types of measurement errors.

SIMULATIONS

In this simulation, we used the above dose model to generate the true dose, assuming every subject received a nonzero exposure. Two scenarios were considered: (1) the exposure effect exists and $\beta = 0.02$ per 10 mSv; (2) no exposure effect, i.e., $\beta = 0$.

The baseline hazard function is assumed to be dependent on the individual's attained age with a log of relative risk of 0.089 per year increase, as obtained from the ORNL study (Xue et al. 2004). The age at entry is simulated from a uniform distribution ranging between 25 and 40, comparable to the range in the ORNL study. For simplicity, we assume each subject has 20 years of radiation exposure and there were no other covariates except the exposure variable. Then, based on the true dose level and an assumption of a 5-y lag, we generate survival times based on a Cox model for each scenario. The number of subjects is chosen to be $N = 200$. A small sample size is chosen here so that it will be computationally feasible to run many iterations. Because of the small size of the simulated data, every subject is followed until death so that the study has 83.6% power to detect an effect of $\beta = 0.02$ (based on simulation results).

Within each scenario, three levels of lognormal random measurement errors were generated: small, medium, and high. We used the uncertainty factor K to represent the magnitude of random errors where $\sigma_\epsilon = logK/1.96$ and K is set to be 1.5, 2.0, or 4.0, respectively. The interpretation of K is the following: 95% of the observed doses measured with random errors, Z, will fall between Y/K to KY where Y is the true dose. For example, if $Y = 1.0$ mSv, then 95% of the observed doses will fall between 0.67 and 1.5 mSv when $K = 1.5$, between 0.5 and 2.0 mSv when $K = 2.0$, or between 0.25 and 4.0 mSv when $K = 4.0$. These three levels correspond to the uncertainty parameter considered in Gilbert (1998), K_3, approximately equal to 0.2, 0.4, and 1.0, respectively; while Gilbert considered the cases with K_3 varied from 0.02 to 1.0. The result suggested that the random error does not influence the risk estimator until K_3 is at least 0.2. Therefore, the uncertainty levels considered in this paper should cover a majority of plausible values that have a non-ignorable impact on risk estimates. After adding the random errors to the true exposure, we then applied the MDL accordingly. Doses below the MDL were set to be zero. The observed doses were used to obtain the estimated β.

The simulation was repeated 500 times. We summarized the simulation results for scenario (1) by calculating the relative bias, defined as the mean of the estimated β relative to the true β, and the mean of the estimated standard error of β based on the observed dose relative to that based on the true dose. We also calculated the proportions of the nominal 95% confidence intervals of β not containing zero and containing the true β. The first proportion is interpreted as the power of the test of a null exposure effect and the second proportion is interpreted as the coverage probability of the nominal 95% confidence interval. For the second scenario when the exposure effect is null, we calculated the average of the estimated β and the proportion of the 95% confidence intervals containing zero. This proportion is interpreted as the coverage probability of the nominal 95% confidence interval as well as one minus the type I error associated with the test of a null exposure effect.

Simulations were performed using S-Plus (S-Plus 6.1, 2002, Insightful Corp.).

With an exposure effect (i.e., $\beta > 0$)

We first sampled the mean parameter μ_i (the mean of the log of the annual dose for subject 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 sampled y_{ij} for $j = 1, \cdots, m$ from a lognormal distribution: $logy_{ij} \sim Normal(\mu_i, 0.4)$ such that each subject has its own mean parameter for his annual exposure. Marginally, $EY_{ij} = 0.545$ mSv, $VarY_{ij} = 0.8013$ mSv^2 and $Corr(logY_{ij}, logY_{ij'}) = 0.6$ for $j \neq j'$. The parameters were chosen so that the distribution of the simulated doses is close to that of the ORNL data we used as an example. Random measurement errors were generated and added on to the true dose.

We let the MDL vary from 0, 0.33, and 0.65 mSv so that the chances of missed doses are 0%, 50%, and 75% on average, respectively. The relative bias was plotted against the uncertainty factor K in Fig. 1a. The plot suggests that with the random measurement error and with or without the error due to MDL, the estimated β is

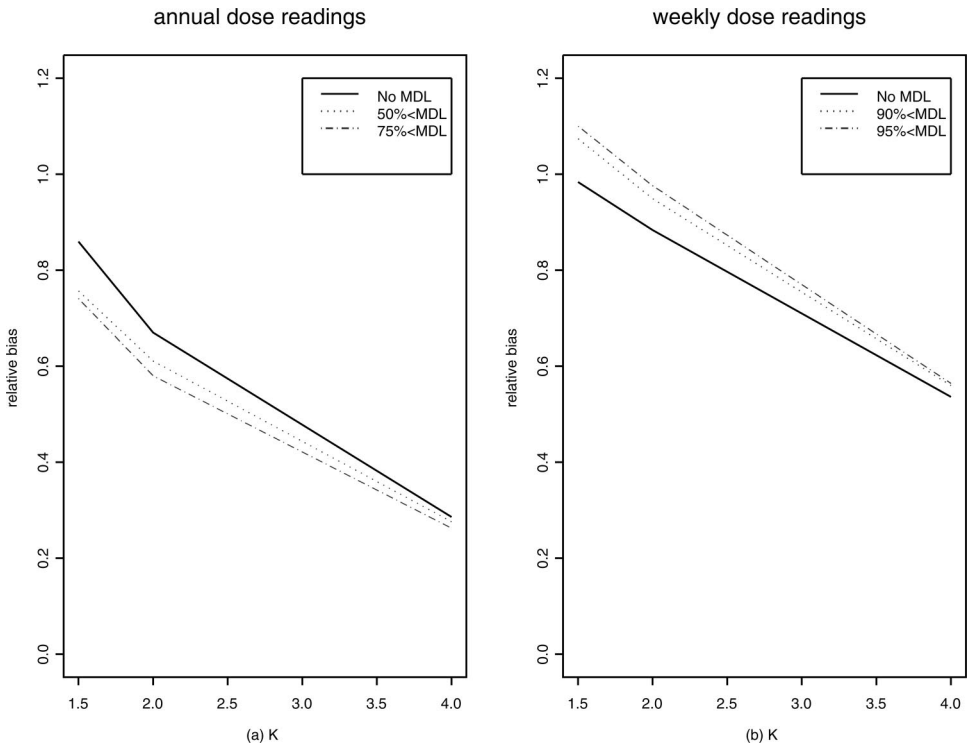

Fig. 1. Relative bias in estimates of β when the exposure is subject to random measurement error and error due to MDL.

biased towards null (the relative bias is less than 1). As the degree of random error goes up, the amount of bias goes up; the amount of bias also goes up as the error due to MDL increases. However, the random error has a stronger effect on the estimate of the relative risk compared to the MDL. The standard errors are also biased downward but to a lesser degree (results not shown).

To see how the errors affect the estimate of β when doses were measured weekly, the weekly doses y_{ijk} for $l = 1, \cdots, 50$; $j = 1, \cdots, n_i$ and $i = 1, \cdots, N$ were sampled from a lognormal distribution such that the annual dose

$$\sum_{l=1}^{50} y_{ijk}$$

approximately follows lognormal(μ_i, σ^2). The parameters were set to be the same as above. A series of weekly readings was added to form the estimate of the annual dose. Weekly multiplicative random measurement errors were generated using the same values of K as the annual readings. The sum of the observed weekly doses is a better estimate of the true annual dose than the observed annual dose when the film badge was monitored annually. This occurs because the random errors from multiple measurements tend to cancel each other out. Thus, the impact of random measurement error is smaller comparing to the annual monitoring case. The MDLs for weekly doses were chosen so that the chances of missed doses were 0% and 90% and 95% on average, respectively. With $\beta = 0.02$, as shown in Fig. 1b, with only random error, when the level of random error is low, there is little bias in the estimate of the log of the relative risk. When the level of random error becomes medium or high, the log of the relative risk is biased downward by 12% and 46%, respectively. When both the random measurement error and censoring due to minimum detection level exist, even with a very high proportion of BMDL censoring (from 90% to 95%), the bias remains smaller compared to that for the annual readings.

With a small level of random error, as indicated in Fig. 1b, the log of the relative risk is overestimated from 7% to 10% as the censoring proportion changes from 90% to 95%. With a medium or a high level of random error, β is underestimated because the underestimation from random error fortuitously offsets the overestimation from censoring due to BMDL.

The results for quarterly readings are similar and therefore not discussed here. We varied the true β from 0.02 to 0.01 and the lag time from 5 to 10 y. Results are all similar and therefore not presented here.

We also used simulations to examine the impact of errors on the inferences of β. Table 1 summarizes the power of the test of $\beta = 0$ when doses were read annually.

As we mentioned earlier, without any errors, the study has 83.6% power to detect a β of 0.02. With the inclusion of random error, the power decreases. The power also decreases as the censoring due to MDL increases. The impact of the random error on the power is larger than that of the error due to MDL. The coverage probability of the true β is less than 95%, indicating that the 95% confidence interval obtained from the regular Cox model is too narrow. Again, the coverage probability decreases as the level of random error increases and the error due to MDL increases, but the random error has a larger impact on the coverage probability than the error due to MDL does.

With no exposure effect (i.e., $\beta = 0$)

We repeated the simulation with $\beta = 0.0$. When $\beta = 0.0$, Table 2 shows that neither the magnitude of the random error nor the level of censoring due to BMDL lead to a biased estimate of β: the mean of the estimated β is very close to 0. The errors in general lead to an underestimation of the standard error. The coverage probability of the 95% confidence interval is smaller but very close to its nominal level, indicating that the interval estimates are valid and the type I error of a null test is close to 5%.

In summary, the simulations show that the impact of the random measurement error on the estimation and inference of the risk parameter is larger than that of the error due to a minimum detection level. Therefore, it is critical to correct for random measurement error if it exists. The impact of random measurement error on estimating the risk parameter is larger for annual readings than for quarterly and weekly readings because, for the latter, the random measurement errors tend to cancel out when the readings are added together and the sum of weekly readings therefore give a better estimate of the true annual dose than a single annual reading. To see this mathematically, conditioning on μ_i, the annual dose obtained from an annual reading Z_{ij} has the following variance:

$$Var(Z_{ij}) = Y_{ij}^2(e^{2\sigma^2} - e^{\sigma^2}), \quad (4)$$

while the annual dose obtained from weekly readings Z_{ijl}'s has the following variance:

$$Var(\sum_{l=1}^{50} Z_{ijl}) = \sum Y_{ijl}^2(e^{2\sigma^2} - e^{\sigma^2}) \le Y_{ij}^2(e^{2\sigma^2} - e^{\sigma^2}), \quad (5)$$

since

$$\sum Y_{ijl}^2 \le (\sum Y_{ijl})^2 = Y_{ij}^2. \quad (6)$$

Censoring due to BMDL induces an underestimation of risk associated with exposure for annual readings, but it can lead to an overestimation of risk with weekly or quarterly readings. Similar findings were seen in Xue et al. (2004). The rationale is that for weekly or quarterly measurements, the chance that the true exposure is below the MDL is so high that most of the subjects ended up with observed doses zero, regardless of whether the actual weekly or quarterly dose was close to (but below) the MDL or far below the MDL. Only a few high weekly or quarterly measurements ended up being observed. Therefore, the difference in observed cumulative doses between a person with a high exposure and a person with a low cumulative exposure was most likely artificially decreased so that the risk associated with the exposure was overestimated. On the other hand, for annual measurements, since the chance of BMDL doses is much lower compared to weekly or quarterly measurements, the observed doses were underestimated but not as much (i.e., the majority of them were still above zeroes) with the higher exposure being underestimated less and the low exposure being underestimated more. Therefore, the difference in observed doses between the high exposure and low exposure was artificially increased so that the

Table 1. Properties of statistical inference on β when doses in annual readings are measured with multiplicative random measurement error and error due to minimum detection level (result obtained with $N = 200$ subjects, 500 simulations, $\beta = 0.02$).

MDL (% of below MDL doses)	Uncertainty factor	% of 95% CI does not include 0 (power) (%)	% of 95% CI include true beta (%)
0 (No MDL)	1.5	79.0	89.4
	2.0	74.9	79.1
	4.0	61.5	36.1
33 (50%)	1.5	77.2	82.2
	2.0	74.1	69.9
	4.0	61.1	31.3
65 (75%)	1.5	75.9	73.1
	2.0	67.0	69.4
	4.0	25.7	57.0

Table 2. Properties of statistical inference on β when doses in annual readings are measured with multiplicative random measurement error and error due to minimum detection level (result obtained with $N = 200$ subjects, 500 simulations, $\beta = 0.00$).

MDL (% of below MDL doses)	Uncertainty factor	Average of estimated beta	Ratio of estimated se with se based on true dose	% of 95% CI include true beta (%)
0 (No MDL)	1.5	0.0008	0.9612	94.2
	2.0	0.0008	0.9305	93.8
	4.0	0.0007	0.6869	94.6
33 (50%)	1.5	0.0007	0.8481	94.4
	2.0	0.0007	0.8251	94.0
	4.0	0.0007	0.6360	94.2
65 (75%)	1.5	0.0006	1.0035	94.4
	2.0	0.0009	0.8978	94.4
	4.0	0.0006	0.6268	94.8

risk associated with exposure was underestimated. The random measurement error on exposure in general leads to an underestimation of the risk estimator. With both errors acting simultaneously, the direction of the bias depends on which error plays a dominant role.

Further, even though the estimate of the risk effect is biased and the power of the null test is reduced by random error and error due to BMDL, the simulation shows that the type I error rate is not affected.

Application to the ORNL study

The mortality experience of employees at the Oak Ridge National Laboratory has been studied extensively in relation to radiation exposures. Checkoway et al. (1985) studied the mortality of 8,375 white males hired at ORNL between the opening of the facility in January, 1943, and the end of 1972 with follow up through 1977, and Wing et al. (1991) extended this follow-up through 1984. They included men who had worked for at least 30 d and for whom there was no record indicating they had been employed at another Oak Ridge facility. In the present analysis, we considered a subset of 3,960 workers who entered the laboratory after 1945 and had worked for at least 3 y and were followed till 1984. The purpose of only considering workers who had worked for at least 3 y is to ensure adequate exposure for all the subjects so that we could evaluate the exposure measurement error models effectively. This results in 550 deaths in a total of 106,840 person-years.

An individual's recorded radiation dose at ORNL was based on film badges from 1944 to 1975 and thermoluminescent dosimeters from 1975 on. Film badges were evaluated weekly from July 1944 to July 1956, when quarterly monitoring was initiated. The minimum detection level was from 0.10 mSv to 0.30 mSv (Morgan 1962; Kerr 1994). Annual monitoring was initiated in 1975 using thermoluminescent dosimeters (MDL = 0.20 mSv). The doses below the minimum detection levels were set to zero. However, not all the observed zero doses are BMDL doses. Some are legitimate zero doses since office workers tended to have no occupational external radiation exposure (Watkins et al. 1997). Among those positive doses, it was estimated that about 98% of the weekly doses were BMDL doses, about 75% of the annual doses were censored due to BMDL (Xue and Shore 2003) and about 80% of the quarterly doses were BMDL doses. Without any adjustment of the missed doses, the average person-year cumulative dose is 4.58 mSv with an assumption of a 20-y lag.

In Xue et al. (2004), the data set described above was used to estimate the relative risk for all-cause mortality, with the exposure adjusted for missed doses due to BMDL using a Monte Carlo method. A log linear Cox proportional hazards model was used. Data on smoking, chemical exposures, and medical exposures to ionizing radiation were not available. The Cox model controlled for the sociodemographic variables of 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 being at a higher socioeconomic level. Active worker status was considered because workers who continued employment, and consequently exposure, tended to be healthier. A lag of 20 y was used. Under a Cox model controlling for sociodemographic factors, the relative risk was estimated to be 1.018 [95% CI = (0.996, 1.040)] per 10 mSv increase in radiation exposure using the observed doses compared to 1.017 [95% CI = (0.997, 1.037)] using the doses adjusted for censoring due to BMDL. This provides no clear indication of a dose-response effect with respect to all-cause mortality. However, another possibility is that there is a dose effect that was masked by exposure measurement errors.

Several papers studied or addressed the issue of missing doses due to BMDL for the ORNL study (Kerr

1994; Wing et al. 1994; Watkins et al. 1997; Frome et al. 1997; Mitchell et al. 1997). However, the magnitude of random measurement error and other types of systematic measurement error have not been examined for the ORNL study. In this paper we evaluate the dose response relationship for the ORNL study taking into account concomitantly both error caused by BMDL and random measurement error.

For the reasons stated earlier, we used simulations to evaluate the effect of random error in addition to the error due to BMDL on risk estimates. First, we generated weekly doses for 1944 to 1956, quarterly doses for 1957 to 1974 and then annual doses for 1975 to 1984, using the dose distribution parameters described in the above simulation study. These parameters were chosen so that the simulated dose would have the same distribution as the annual exposure in the ORNL study. To ensure that the simulated subjects would have the same employment history, age distribution, and censoring time distribution as the ORNL data set, we bootstrapped (sample with replacement) individuals from the ORNL subset using their hire year, entry age, job termination year, and censoring time to generate our simulated data set. Since some of the annual doses in the ORNL study were legitimate zero doses because they worked in a job with no potential for exposure, we also obtained the indication of the true zero doses from the bootstrap sample and set those doses to be zero in our simulated dose data. We used the hire year and the job termination year for each individual in the bootstrap sample to define the cumulative exposure for this subject and his entry age and attained age over time to define his baseline hazard rate. His survival time was generated based on the hazard function of $\lambda_0 exp[\beta Y(t-20) + \gamma X(t)]$ where $\beta = 0.018$ and $\gamma = 0.089$ were obtained from Xue et al. (2004) and $X(t)$ is the attained age. Since the average entry age was 30 y, we set $\lambda_0 = 1/40$, implying that the expected survival time since the entry into the study for a subject without any radiation exposure was 40 y. If a subject's survival time was more than his censoring time obtained from the bootstrap sample, then his survival time was censored at his censoring time. Several levels of the random measurement errors described in the previous section were generated and added only to the true positive doses. The MDL levels were then applied to the doses measured with measurement error, and doses below the MDL were treated as zeroes. The relative risk was estimated based on the doses with random measurement error and error due to MDL. The simulation was repeated 100 times. Due to computer intensiveness, we set the size of the simulated data to be 1,320, a third of the size of the ORNL subset. Table 3 summarizes the results.

Table 3 indicates that when the magnitude of the random errors is small or medium, the estimate of β is very close to being unbiased. As the magnitude of random errors increases to a high level, the relative bias increases to 44%. There is a large uncertainty association with the bias factors. If the true relative risk is 1.018, then 95% of the time we will get a relative risk estimate in [1.013($= e^{0.018\times0.7406}$), 1.040($= e^{0.018\times2.2391}$)] when the random error is small; (1.011, 1.039) with a moderate random error; and (1.004, 1.027) with a large random error. Nevertheless, the confidence interval associated with the estimated log of the relative risk has 100% chance to exclude 0, indicating that the study without correcting for exposure measurement errors has 100% power to detect a significant β. This is an important finding. Measurement error in general reduces power, but with the size of the study and the magnitude of the effect, the levels of random error and error due to BMDL we considered here do not affect the power.

To see if the result differs when the true β is smaller, we set $\beta = 0.009$ as well as $\beta = 0.003$, a half and a sixth of the size of the previous β, respectively, and repeated the simulations. The conclusions generally remain the same therefore the results are not presented here.

We also used the simulation study to evaluate the type I error of the null test: $\beta = 0$ by setting $\beta = 0$ and re-running the simulation. Table 4 shows that, consistent with the findings in the previous section, 95% of the estimated β are close to 0 and the estimated 95% confidence interval is close to its nominal value. This suggests that without correcting for the two measurement errors, the study maintains its type I error.

In summary, the hypothesis testing is minimally affected by the levels of errors considered in the study. If

Table 3. Estimating β under various levels of multiplicative random measurement error with the same employment history, age, follow-up time, monitoring frequency, and censoring due to the minimum detection level as the ORNL data set and $N = 1{,}320$ and $\beta = 0.018$.

Uncertainty factor	Ratio of estimated beta with true beta	95% CI for the ratio	% of 95% CI does not include 0 (power) (%)	% of 95% CI include true beta (%)
1.5	1.0341	(0.7406, 2.2391)	100	11.1
2.0	0.9864	(0.6202, 2.1677)	100	0.00
4.0	0.5682	(0.2492, 1.4936)	100	3.37

Table 4. Estimating β under various levels of multiplicative random measurement error with the same size, follow-up time, monitoring frequency, and censoring due to the minimum detection level as the ORNL data set and $\beta = 0.00$.

Uncertainty factor	Average of estimated beta	95% empirical CI for beta	% of 95% CI includes 0 (%)
1.5	0.0	(−0.0006, 0.0005)	96.0
2.0	0.0	(−0.0007, 0.0006)	93.0
4.0	0.0	(−0.0004, 0.0004)	98.0

the true β is of a magnitude of at least 0.003 per 10 mSv (a minimal level of clinical importance), with the same study population and a size of at least 1,320 subjects, without correcting for measurement errors, our simulation studies show that we would still have almost 100% chance to detect a significant β. The fact that the naive estimate of β was not significant reinforces our conclusion that radiation exposure is not significantly associated with total mortality in this study. Since the uncertainty factor in most of the occupational radiation studies in the U.S. rarely exceeded the levels we considered in the simulations (Gilbert 1998), it was then concluded that random measurement error is not likely to be a major concern in the ORNL study.

DISCUSSION

Measurement error is ubiquitous for occupational radiation dose estimates. New techniques and more advanced dosimeters can reduce the amount of measurement error but can never completely eliminate it. Since measurement errors in dose estimates may distort the results of epidemiologic analyses, it is important to correct for them when estimating dose-response relationships. There are two broad types of measurement error: systematic and random, as discussed in the introduction of this paper. This paper focuses on laboratory random measurement errors and a special type of systematic measurement error. The ascertainment and correction for other types of systematic measurement error requires detailed evaluation of the dosimetry practices for the particular study but statistically is quite straightforward. Laboratory random measurement error is introduced in film badge and dosimeter readings. We assumed that approximately the same relative degree of error was introduced every time a film badge or a dosimeter was read. Therefore, the estimated annual dose resulting from 50 weekly readings tends to be more stable than one annual reading since the error associated with weekly readings are "smoothed out" over time. However, there may be other sources of random error such as environmental error that depends on the frequency of readings: the longer the period the dose gets accumulated before reading, the more stable the estimate is, the less the error. Therefore, in such models it is more appropriate to assume that the relative degree of random error decreases with the frequency of readings: i.e., the annual readings have the smallest level of error while the weekly readings have the highest level of error.

Development and implementation of analytical statistical methods to correct for measurement errors are challenging because more than one type of measurement error can be present, and models for errors associated with cumulative exposure are complex. Thus, we proposed one should make use of simulation studies to evaluate empirically the impact of random measurement error and error due to BMDL on relative risk estimation under various assumptions. One can then make statistical inference on the risk parameter correcting for random errors and errors due to BMDL. One important finding from the simulation studies is that the random error has a larger impact on the relative risk estimates compared to the error due to MDL. Therefore, it is essential to assess the impact of random error and correct for it appropriately if necessary. Another important finding from the simulation studies is that for the size of the ORNL study and the plausible levels of random errors, the random error and error due to BMDL do not affect the properties of the hypothesis test of a null β: the power remains high and the type I error is close to the nominal level. Therefore, any conclusion drawn from the "naive" hypothesis test is still valid. From the results based on the ORNL study, we are able to infer that there is no significant relationship between radiation exposure and mortality, although there is still some uncertainty because of not having information on certain covariates such as smoking.

Simulation studies are flexible, permitting a range of assumptions about the magnitude of random errors and the proportion of censoring due to MDL. The models of measurement error can also be changed to permit additive measurement error or Berksonian measurement error and various correlation structures among the measurements. Extension of the simulation study to allow various types of systematic measurement error is straightforward. For the situations where the analytical approach is not available, simulations provide a useful alternative. Finally, our simulation method can be used rather generally for a wide variety of types of cumulative exposure data.

Acknowledgments—This work was supported by NIOSH/CDC grant 1R01 CCR215746 and by Center Grants ES-00260 from the NIEHS and CA016087 from the NCI to New York University School of Medicine.

REFERENCES

Armstrong BG. The effects of measurement errors on relative risk regressions. Am J Epidemiol 132:1176–1184; 1990.

Cardis E, Vrijheid M, Blettner M, Gilbert E, Hakama M, Hill C, Howe G, Kaldor J, Muirhead CR, Schubauer-Berigan M, Yoshimura T, Bermann F, Cowper G, Fix J, Hacker C. Risk of cancer after low doses of ionizing radiation—retrospective cohort study in 15 countries. Br Med J 331:77–81; 2005.

Checkoway H, Mathew RM, Shy CM, Watson JE, Tankersley WG, Wolf SH, Smith JC, Fry SA. Radiation, work experience, and cause-specific mortality among workers at any energy research laboratory. Br J Indust Med 42:525–533; 1985.

Clayton DG. Models for the analysis of cohort and case-control studies with inaccurately measured exposures. In: Dwyer JH, Lippert M, Feinleib M, Hoffmeister H, eds. Statistical models for longitudinal studies of health. New York: Oxford University Press; 1992: 301–331.

Cochran WG. Errors of measurement in statistics. Technomet 10:637–665; 1968.

Comprehensive Epidemiologic Data Resource. Washington, DC: U.S. Department of Energy Office of Epidemiologic Studies; 20585; 1999.

Dupree-Ellis E, Watkins J, Ingle JN, Phillips J. External radiation exposure and mortality in a cohort of uranium processing workers. Am J Epidemiol 152:91–95; 2000.

Frome EL, Cragle DL, Watkins JP, Wing S, Shy CM, Tankersley WG, West CM. A mortality study of employees of the nuclear industry in Oak Ridge, Tennessee. Radiat Res 148:64–80; 1997.

Gilbert ES. Accounting for errors in dose estimates used in studies of workers exposed to external radiation. Health Phys 74:22–29; 1998.

Gilbert ES, Fix JJ. Laboratory measurement error in external dose estimates and its effects on dose-response analyses of Hanford worker mortality data. Richland, WA: Pacific Northwest Laboratory; PNL-11289; 1996.

Gilbert ES, Omohundro E, Buchanan J, Holter N. Mortality of workers at the Hanford site: 1945–1986. Health Phys 64:577–590; 1993.

Gilbert ES, Fix JJ, Baumgartner WV. An assessment of bias and uncertainty in recorded does from external sources of radiation for workers at the Hanford site. Health Phys 70:336–345; 1996.

Gilliland FD, Hunt WC, Archer VE, Saccomanno G. Radon progeny exposure and lung cancer risk among non-smoking uranium miners. Health Phys 79:365–372; 2000.

Kerr GD. Missing dose from mortality studies of radiation effects among workers at Oak Ridge National Laboratory. Health Phys 66:206–208; 1994.

Mitchell TJ, Ostrouchov G, Frome EL, Kerr GD. A method for estimating occupational radiation dose to individuals, using weekly dosimetry data. Radiat Res 147:195–207; 1997.

Morgan KZ. Techniques of personnel monitoring and radiation surveying. In: Nuclear instruments and their uses. New York: Wiley; 1962: 391–469.

Prentice RL. Covariate measurement errors and parameter estimation in a failure time regression model. Biometrika 69:331–342; 1982.

Richardson D, Ciampi A. Effects of exposure measurement error when an exposure variable is constrained by a lower limit. Am J Epidemiol 157:355–363; 2003.

Ritz B. Radiation exposure and cancer mortality in uranium processing workers. Epidemiol 10:531–538; 1999.

Thomas D, Stram D, Dwyer J. Exposure measurement error: influence on exposure-disease. Relationships and methods of correction. Annu Rev Publ Health 14:69–93; 1993.

Watkins JP, Cragle DL, Frome EL, West CM, Crawford-Brown DJ, Tankersley WG. Collection, validation, and treatment of data for a mortality study of nuclear industry workers. Appl Occup Environ Hyg 12:195–205; 1997.

Wiggs LD, Johnson E, Cox-DeVore C, Voelz G. Mortality through 1990 among white male workers at the Los Alamos National Laboratory: considering exposures to plutonium and external ionizing radiation. Health Phys 67:577–588; 1994.

Wing S, Shy C, Wood J, Wolf S, Cragle D, Frome E. 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.

Wing S, West CM, Wood JL, Tankersley W. Recording of external radiation exposures at Oak Ridge National Laboratory: implications for epidemiological studies. J Exp Analy Environ Epidemiol 4:83–93; 1994.

Xue X, Shore RE. Methods for estimating occupational radiation doses subject to minimum detection levels. Health Phys 84:61–71; 2003.

Xue X, Shore RE, Ye X, Kim MY. Estimating the dose response relationship for occupational radiation exposure measured with minimum detection level. Health Phys 87:397–404; 2004.

■ ■