RADIATION RESEARCH 123, 275–284 (1990)

Allowing for Random Errors in Radiation Dose Estimates for the Atomic Bomb Survivor Data

DONALD A. PIERCE, DANIEL O. STRAM, AND MICHAEL VAETH

Department of Statistics, Radiation Effects Research Foundation, Hiroshima, Japan

PIERCE, D. A., STRAM, D. O., AND VAETH, M. Allowing for Random Errors in Radiation Dose Estimates for the Atomic Bomb Survivor Data. *Radiat. Res.* 123, 275–284 (1990).

The presence of random errors in the individual radiation dose estimates for the A-bomb survivors causes underestimation of radiation effects in dose–response analyses, and also distorts the shape of dose–response curves. Statistical methods are presented which will adjust for these biases, provided that a valid statistical model for the dose estimation errors is used. Emphasis is on clarifying some rather subtle statistical issues. For most of this development the distinction between radiation dose and exposure is not critical. The proposed methods involve downward adjustment of dose estimates, but this does not imply that the dosimetry system is faulty. Rather, this is a part of the dose–response analysis required to remove biases in the risk estimates. The primary focus of this report is on linear dose–response models, but methods for linear-quadratic models are also considered briefly. Some plausible models for the dose estimation errors are considered, which have typical errors in a range of 30–40% of the true values, and sensitivity analysis of the resulting bias corrections is provided. It is found that for these error models the resulting estimates of excess cancer risk based on linear models are about 6–17% greater than estimates that make no allowance for dose estimation errors. This increase in risk estimates is reduced to about 4–11% if, as has often been done recently, survivors with dose estimates above 4 Gy are eliminated from the analysis.

INTRODUCTION

Much work at the Radiation Effects Research Foundation (RERF) involves estimation of the A-bomb survivors' radiation dose response for various end points, using estimates of individual radiation doses that have substantial uncertainties. Random errors in dose estimates cause systematic biases in estimates of risk based on a linear dose–response model, and also distort the shape of the dose response. In this paper we develop statistical methods for dealing with these problems, with emphasis on clarifying some of the issues involved. The primary motivation for this work was analysis of cancer data, but some attention is given to analysis of data for chromosomal aberrations to illustrate how the methods depend on the nature of the response data. The focus is on fitting linear dose–response models, but brief consideration is given to the extension required for linear-quadratic models. In another report Pierce and Vaeth (*1*) use that extension to study the shape of the cancer dose–response curve.

The problem of errors in doses for RERF data has been investigated in some detail by Jablon (*2*) and Gilbert (*3*). Jablon's report is particularly useful because of his analysis of the probable form and magnitude of errors in dose estimates. The general conclusions of the Gilbert paper are quite similar to those presented here, but the details of our development and proposed implementation of statistical methods are substantially different.

A generally useful classical reference for the "errors in covariables" problem is Cochran (*4*), and an entry into the recent literature is provided by papers in a recent workshop (*5*). A textbook with broad coverage is Fuller (*6*), but this book deals little with the particular needs and approach here. More closely related methods are investigated by Armstrong (*7*), Prentice (*8–10*), and Clayton (*11*).

Consideration is given only to what might be called "random" errors in dose estimates, due largely to uncertainties regarding the survivors' location and shielding, in contrast to those of a more systematic nature, such as in the yields of the bombs. No attempt is made here to investigate the actual nature of the dose-estimation errors. Some preliminary discussion of precision in the Dosimetry System 1986 (DS86) estimates, as well as a useful summary of the entire system, is given in Chapter 9 of the DS86 Report (*12*). The assessment of errors given by Jablon (*2*) for the previous dosimetry system remains relevant, because it focused on errors due to uncertainties about survivor location and shielding, and the same basic input is used in the DS86. Further research is needed on both the required statistical methods, and the ascertainment of the nature and magnitude of errors in the DS86 estimates.

For much of the paper the term dose will refer to "in air" tissue kerma at the location of the survivor corrected for shielding and terrain. Adjustments for errors in the tissue kerma estimates will be derived, resulting in factors which can be applied to organ doses. The alternative of developing more specific adjustments for doses to various organs seems

0033-7587/90 $3.00

needlessly complicated. The primary weakness in basing adjustments on tissue kerma is that organ dose estimates generally have somewhat larger errors than do tissue kerma estimates, and the magnitude of the additional errors may depend on the organ considered. These points can be seen from the observation that the error in an organ dose estimate depends, to an extent depending on the organ under consideration, on uncertainty about which direction the survivor was facing when exposed.

BASIC STATISTICAL ISSUES

Suppose the true dose response for some outcome y is of linear form $\alpha + \beta x$, where x is the *true dose* and α, β are parameters to be estimated. Then due to the linearity in x, the expected response among those with *estimated dose* z is $\alpha + \beta\ \mathrm{Avg}(x|z)$, where $\mathrm{Avg}(x|z)$ is the average true dose among those with estimated dose z. The numerical values of the parameters α and β are the same in these two expressions. It is useful to think of $\mathrm{Avg}(x|z)$ as the average x-value for those with *approximately* the same z-value. Of course $\mathrm{Avg}(x|z)$ is not directly available from the data, but consideration of this quantity, as a function of z, is the key to understanding how biases arise when the relationship of y to z is used to draw inferences about the relationship of y to x, and how to approach correcting this problem.

This argument depends on the dose response being linear in x, and the focus here is on that case. However, it can be generalized in various ways; extension to linear-quadratic models is discussed briefly. It is noted that if the nonlinearity in the expected value of y given x is not very great in the *restricted* range of x-values consistent with an estimated dose z, then a graph of y versus $\mathrm{Avg}(x|z)$ would still reflect the true shape of the dose response. Corrections can be made to the approximation in this, but the point remains that the relationship between z and $\mathrm{Avg}(x|z)$ is of general importance.

It is easy to confuse the notion of $\mathrm{Avg}(x|z)$ with that of $\mathrm{Avg}(z|x)$. Lack of bias in the dosimetry system should be taken to mean that $\mathrm{Avg}(z|x) = x$. It will be seen that even when the dose estimates are unbiased in this sense, $\mathrm{Avg}(x|z)$ is not equal to z. The hypothetical example developed in Table I may clarify this. Rounding to the nearest gray, the numbers of survivors in the true dose categories there are roughly those of the RERF Life Span Study (LSS) cohort. Consider a highly artificial statistical model for errors, such that with probability 0.5 the estimate is correct, with a probability of 0.25 of a 1-Gy error in either direction. The table indicates the expected numbers of survivors in a cross-classification of true and estimated doses. Consideration of the *rows* of the table shows that, for each true dose, the average estimated dose is equal to that true dose and thus the estimates are unbiased in the sense of $\mathrm{Avg}(z|x) = x$. On the other hand, $\mathrm{Avg}(x|z)$ is computed by averaging

TABLE I
Artificial Example Indicating Basic Concepts

True dose x (Gy)	Estimated dose z (Gy)						Number of survivors
	1	2	3	4	5	6	
—	—						—
—	—	—					—
2	250	500	250				1000
3		75	150	75			300
4			33	66	33		132
5				15	30	15	60
—					—	—	—
$\mathrm{Avg}(x\|z)$	—	—	2.50	3.62	—	—	

true dose according to the frequencies in each *column;* some selected values are given at the bottom of the table. The average true dose for those in an estimated dose category is less than the estimated dose. This is seen to result from there being far more survivors in each column who are 1 Gy below the estimate than 1 Gy above the estimate. There is a skewness in each column which is, in turn, due to the skewness of the marginal distribution of true doses.

The error distribution in Table I was chosen largely for simplicity, but it illustrates some important points. It is often thought that two major aspects of the problem resulting from errors in doses are that the errors are: (i) greater for larger doses, and (ii) rather symmetric on a logarithmic scale, resulting in larger overestimates than underestimates. These two factors are relevant, but it can be seen from Table I, which has neither of these features, that the fundamental problem would exist even without them.

The achievement of $\mathrm{Avg}(z|x) = x$, for all values of x, is a reasonable aim of a dosimetry system. It is useful to consider this issue in terms of an idealized dosimetry system which would give exact estimates were it not for errors in input variables such as the individual's location, shielding, etc. Then unbiased estimation of these input variables would lead, to an approximation which might be improved by transforming to a log scale, to $\mathrm{Avg}(z|x) = x$. We would say that dose estimates from such an idealized system are unbiased. Whether the DS86 estimates are in fact unbiased in this sense is not considered in this paper; we will assume that they are unbiased, on a log scale.

However, this will not ordinarily imply lack of bias in parameter estimates for dose–response models. The bias in linear risk estimation based on estimated dose z is related directly to the extent that $\mathrm{Avg}(x|z)$ differs from z. This is due to the fact noted above that, when the expected response is given by $\alpha + \beta\ x$, the expected response among survivors with a given value of z is $\alpha + \beta\ \mathrm{Avg}(x|z)$. It will be shown here that for the LSS cohort $\mathrm{Avg}(x|z)$ is substantially less than z in the higher range of estimated dose.

The statistical methods suggested for linear risk models involve the replacement of the estimated doses z by adjusted estimates $\text{Avg}(x|z)$, prior to dose–response analyses. True values of $\text{Avg}(x|z)$ are not known, but these may be approximated as explained here. For some types of data additional modifications are required to allow for the variation introduced by errors in dose estimates; for the cancer data this is generally not necessary. Additional, or sometimes different, adjustments are required for nonlinear models.

It is emphasized that this adjustment of estimated doses is not intended to correct any fault in the dosimetry system. When the distribution of true doses is very skewed, the two conditions $\text{Avg}(z|x) = x$ for all x and $\text{Avg}(x|z) = z$ for all z are simply incompatible. Coping with the fact that $\text{Avg}(x|z)$ is less than z for large dose estimates falls more within the realm of dose–response analysis than that of the dosimetry system.

STATISTICAL MODELS AND METHODS

We first discuss statistical models for errors and a method for estimation of $\text{Avg}(x|z)$. After that, we indicate why and how adjusted estimates should be used in dose–response analyses, and give further details of the statistical procedures involved.

Estimation of Avg(x|z)

$\text{Avg}(x|z)$ is the mean of the conditional distribution of x among cohort members having (approximately) the same value of z; this conditional distribution is written as $f(x|z)$. Other quantities computed from $f(x|z)$, such as $\text{Avg}(x^2|z)$ and the standard deviation $\text{SD}(x|z)$ will also be of interest.

The density function $f(x|z)$ can be obtained from the relationship

$$f(x|z) \propto f(x)f(z|x), \tag{1}$$

where $f(x)$ is the marginal distribution of true doses and $f(z|x)$ is the distribution of the estimated doses for a given true dose. The proportionality in Eq. (1) is with respect to x. Several assumed models for $f(z|x)$ will be used, to explore the sensitivity of final results to this choice. Numerical integration is used here to determine moments of $f(x|z)$ from Eq. (1), in order to avoid artificial constraints due to mathematically convenient choices of models. An estimate of the distribution $f(x)$ of true doses in the study population must be made; this will be done by adjusting the observed distribution of estimated doses z.

Before turning to details of these points, we discuss the intended meaning of $f(x|z)$, since these distributions are central to the approach. The intention is not to consider a particular survivor's true dose as a random variable, but rather to think of the distribution $f(x|z)$ as representing the actual, but unobserved, distribution of true doses in the cohort among those with (approximately) a given estimated dose z. Considering Table I may clarify this. From this viewpoint the distribution $f(x)$ in Eq. (1) represents the actual distribution of true doses in the cohort, rather than that of a population from which the cohort is considered to be a sample.

Careful definition of the meaning of $f(x|z)$ involves subtle statistical issues, and the point is raised here for two reasons. Replacing carefully obtained dose estimates z by adjusted values $\text{Avg}(x|z)$ in dose–response analyses certainly deserves careful thought, and may be controversial. The authors felt that it might be considered inappropriate to treat a given survivor's true dose as a random variable. Further, in the general statistical literature on errors in covariables a major distinction is made between considering the true values x as random variables or as fixed unknown quantities. In applications, however, this distinction is unclear.

The meaning of $f(x|z)$ may be clearest for grouped data, but there are two reasons why it is helpful to carry out the development here without such grouping. Numerical results here can be used for whatever grouping on estimated exposures may be appropriate for a given investigation, and even for ungrouped data when such analysis is feasible. Further, in the second part of this section it will be important to consider the distribution of x given z for individual survivors. In concrete terms it may be best to think of this as representing the x-value for a survivor selected at random from those with approximately a given estimated dose z.

Attention is now given to choice of $f(x)$, the distribution of true doses in the cohort. Because the distribution of survivors' locations relative to hypocenters of the two bombs differed markedly, $f(x)$ differs between cities (*2*). The choice of $f(x)$ and the computation of $\text{Avg}(x|z)$ is therefore done by city. The following procedure adjusts the observed distribution of z for each city to arrive at estimates of the distribution of x.

For each city the proportion of those with dose estimates greater than any given value of z can be fitted very well by expressions of the form $\exp(-\theta_1 z^{\theta_2})$, which is known as the Weibull model. (The values of θ_2 are less than one, corresponding to a monotonically decreasing density function.) It seems adequate to assume the same parametric form for the distribution of x, and estimate the parameters θ_1 and θ_2. This is done here by choosing these parameters so that, with a given model for the errors, the theoretical distribution of the z's induced by $f(x)$ and $f(z|x)$ agrees well with the observed distribution of z's. In particular this is done by computing this induced distribution numerically, for given values of θ_1 and θ_2 and choice of an error model, and using a direct search to find suitable parameter values for $f(x)$. Since this depends on the choice of an error model $f(z|x)$,

the results are given after discussing the error models considered here.

The lognormal model is of primary interest for $f(z|x)$, where $\log(z)$ is assumed to be normally distributed with mean $\log(x)$ and standard deviation independent of x. Two other models are considered to explore the effects of this assumption. The lognormal model was suggested by Jablon (*2*), after consideration of the nature of the major sources of error. A key feature is that on the original scale the magnitude of errors is proportional to the level of true dose; this feature seems essential to any model for $f(z|x)$. For the range of values used here, the standard deviation (SD) of $\log(z)$ is approximately equal to the coefficient of variation (CV) of z; that is, standard deviation of z relative to its mean. The precise relationship for lognormal models is that $\mathrm{CV}(z)^2 = \exp\{\mathrm{SD}[\log(z)]^2\} - 1$. For example, values of 0.30, 0.35, and 0.40 for $\mathrm{SD}[\log(z)]$ correspond to 30.7%, 36.1%, and 41.7% for $\mathrm{CV}(z)$. For simplicity, such a model with $\mathrm{SD}(z) = 0.30$ will be referred to as "lognormal 30% error," and so forth.

Jablon concluded that the lognormal 30% error model was worth particular attention, but that somewhat larger errors should also be seriously considered. Results are given for the lognormal 30% and 40% error models and for two types of models which are not lognormal, but which have $\mathrm{CV}(z) = 40\%$. Some final numerical results are also given for the lognormal 35% error model.

The most serious concern with the lognormal model is not the assumption of symmetry on the log scale, but rather the more specific assumption of normality. The normal distribution is noteworthy, indeed notorious, for having very "light tails"; that is, it allows for few deviations which are large in relation to the size of more frequent ones. Since results here might be sensitive to the assumed chance of extreme errors, investigating a distribution with "heavier tails" is important. A natural choice is a contaminated lognormal distribution, taking $\log(z)$ as having an 85% chance of being normally distributed with standard deviation 0.30, and a 15% chance of being normally distributed with standard deviation 0.75. These parameters were chosen to give another 40% error model for z, but with a quite different shape for the distribution.

A fourth model was chosen to examine departures from the assumption of symmetry on the log scale, with consequent skewness on the original scale. This model takes z as normally distributed, with mean x and standard deviation proportional to x so that $\mathrm{CV}(z) = 40\%$. Although approximate symmetry on a log scale is more natural, this rather extreme departure nevertheless provides useful information about the effect of the shape of the distribution.

In summary, results will be given for four error models: A, lognormal 30% error; B, lognormal 40% error; C, contaminated lognormal 40% error; and D, normal 40% error.

The dependence of the dose-estimation errors on the shielding characteristics and perhaps other factors should be considered. Further information about the DS86 dosimetry system will be needed to evaluate such refinements.

Use of these Adjustments for Dose–Response Analyses

We first consider models in which the expected response is linear in true dose. Some comments on nonlinear models are given at the end of this section. The aim is to develop more carefully the argument of the previous section, suggesting replacement of the z-values by $\mathrm{Avg}(x|z)$ for dose–response analyses. In particular, it is clear that one should not simply think of values of $\mathrm{Avg}(x|z)$ as *equivalent* to the true x's, and a primary aim is to see what additional allowances may be required.

Two specific types of response data, i.e., cancer and chromosomal aberrations, will be considered below; but first consider in general a response variable y following a statistical model, in terms of true doses, of the form

$$y = \alpha + \beta x + \text{error}. \tag{2}$$

Standard maximum likelihood methods for the data of interest can be implemented as weighted least-squares analysis under this model. For data of interest here the variance of "error" depends on x, which calls for weighted regression, and also on α and β, which calls for iterative methods.

For a survivor selected at random from the cohort, with estimated dose z, Eq. (*2*) can be re-expressed as

$$y = \alpha + \beta\,\mathrm{Avg}(x|z) + \beta[x - \mathrm{Avg}(x|z)] + \text{error}. \tag{3}$$

In this setting x is a random variable with distribution $f(x|z)$ so the term $\beta[x - \mathrm{Avg}(x|z)]$ becomes an additional "error term" in the model, with expected value zero and variance $\beta^2\mathrm{Var}(x|z)$. Thus the variance of x among those with a given z becomes an important quantity, which can be computed similarly to $\mathrm{Avg}(x|z)$.

The primary result is therefore that the datum for an individual with estimated dose z can be expressed as

$$y = \alpha + \beta\,\mathrm{Avg}(x|z) + \text{error}^*, \tag{4}$$

where error* is the combination of the two error terms in Eq. (3). In the sense of averaging over *both* sampling variation and errors in dose estimation, the expected value of the error* term is zero, implying that weighted linear regression analysis based on Eq. (4) will yield unbiased estimates. (There is actually a small, ordinarily negligible, bias due to iterative estimation of weights for the regression). Further, it follows from a standard calculation that

$$\text{Var(error*)} = \beta^2 \text{Var}(x|z) + \text{Avg}[\text{Var(error)}], \quad (5)$$

where the last term is the average, in the distribution $f(x|z)$, of the variance of the error in model (2). Thus Eq. (5) provides the proper weights for fitting model (4) by weighted least squares, and also, ultimately, the standard errors of the parameter estimates.

For some applications the only required modification to standard methods is to replace z by $\text{Avg}(x|z)$. In others one must additionally modify the weights for the regression to accommodate dose errors. The distinction between these depends on the relative size of the two terms in Eq. (5).

We consider two primary types of data that are of interest at RERF: (i) cancer incidence or mortality, and (ii) chromosomal aberration prevalence. A more detailed discussion of the following issues is given in an appendix to Pierce *et al.* (*13*).

For cancer analyses the datum on each individual is essentially binary, making the second term (from sampling errors) in Eq. (5) much larger than the first term (from dose errors). Moreover, the term Avg[Var(error)] in Eq. (5) depends upon $\text{Avg}(x|z)$ in the same way that Var(error) depends on x. Because of these two results, the only required change to standard analyses is to replace z by $\text{Avg}(x|z)$. The conclusions just drawn agree with more careful analysis based on statistical models for what are called "survival data".

For chromosomal aberration analyses the datum on each individual is the proportion of about 100 examined cells which exhibit an aberration. The binomial probability model describes the sampling error for such data. Under the binomial model the sampling error is inversely proportional to the number of cells examined, and is thus much smaller than for the cancer data. Calculations show that the two terms in Eq. (5) are roughly of the same magnitude, and therefore allowance must be made for the additional variation in the data y, which is due to dose-estimation errors. However, these data clearly exhibit more dispersion than predicted by the binomial model, and specially developed methods have been used for some time; see, for example, Preston *et al.* (*14*). These methods are based on an empirical model for the variance, which is remarkably similar to Eq. (5), which incidentally lends support to the modeling of this paper. The conclusion is that although binomial methods modified only by replacing z by $\text{Avg}(x|z)$ are not appropriate, the special methods given by Preston *et al.* (*14*) require only that modification.

The above results pertain to models in which the expected response is linear in true dose. Corresponding development for nonlinear models is generally less tractable, but there is a natural extension for the important case of linear-quadratic models, i.e., those where $\alpha + \beta x$ in Eq. (2) is extended to $\alpha + \beta x + \gamma x^2$. In this case unbiased estimation would be achieved by extending the above methods to regression on both $\text{Avg}(x|z)$ and $\text{Avg}(x^2|z)$. This last quantity can be calculated similarly to $\text{Avg}(x|z)$; indeed the need for this has already arisen above in terms of $\text{Var}(x|z)$, which is $\text{Avg}(x^2|z) - [\text{Avg}(x|z)]^2$. Caution should be taken in using this approach, since $\text{Avg}(x^2|z)$ will probably be more sensitive to the choice of error model than is $\text{Avg}(x|z)$. Further details on the effects of dose errors when fitting linear-quadratic models are given by Pierce and Vaeth (*1*).

The development of methods in this paper is for the case of a single covariable, dose, and it is important to consider the effects of additional covariables measured without error. Such covariables can be included in the dose–response analysis without further modifications if they have no effect on the distribution $f(x|z)$. This is most clearly thought of in terms of their potential effect on either $f(x)$ or $f(z|x)$. The primary additional covariables of interest in the LSS are city, sex, and age-at-exposure. The effects of city on $f(x)$ are accounted for here by calculating the adjustments $\text{Avg}(x|z)$ in a city-specific manner. The other primary covariables, sex and age-at-exposure, are unlikely to affect substantially either $f(x)$ or $f(z|x)$.

An important point is that $f(x|z)$ may depend on covariables which are not in the model for the expected value of y. An example of this is shielding category. Although further attention to this point is of interest, it is noted that failing to incorporate such covariables only decreases the *precision* of estimates of the dose response, rather than introducing *biases.* It may be that this loss in precision is small, and that incorporation of shielding category in computing $\text{Avg}(x|z)$ is less important than might first be thought.

Another important issue arises in analysis of response data available only for a *subset* of the cohort, such as that on chromosomal aberrations. If the selection of the subcohort is based substantially on estimated dose, as it has been for study of chromosomal aberrations, then it would be wrong to use the distribution of the z's *in the subcohort* to arrive at estimates of $\text{Avg}(x|z)$. More precisely, it would be wrong to use this distribution in exactly the same way that we use the distribution of z's for the entire cohort in this paper. The same estimates of $\text{Avg}(x|z)$ that we derive here should be used for analysis of the subcohort data. This is because the distribution $f(x|z)$ is not changed by selection on z. If the distribution of z's for a subcohort selected on estimated dose were to be used as in this paper, allowance would have to be made for the effect of the selection on the error model $f(z|x)$. The error model used here is appropriate for the entire LSS cohort. This point also applies to analysis of data for the Adult Health Study subcohort, which is followed up by clinical examinations.

RESULTS

Estimates of the city-specific distributions $f(x)$ of true doses are now given, based on the method explained in the

TABLE II
Numbers of Survivors in LSS DS86 Cohort Exceeding Selected Values of In Air Tissue Kerma

Kerma (Gy)	Hiroshima	Nagasaki
0.1	14632	3420
0.5	5099	1501
1.0	2229	801
1.5	1311	444
2.0	820	282
2.5	532	185
3.0	338	129
3.5	218	100
4.0	171	85
4.5	128	71
5.0	108	55

TABLE III
Adjusted Dose Estimates for the Four Error Models

Estimated kerma (Gy)	Avg($x\|z$) Hiroshima A[a]	B	C	D	Nagasaki A	B	C	D
0.5	0.50	0.50	0.50	0.51	0.51	0.51	0.51	0.53
1.0	0.96	0.94	0.94	0.97	0.98	0.97	0.97	1.00
2.0	1.84	1.73	1.75	1.81	1.89	1.82	1.83	1.88
3.0	2.66	2.45	2.47	2.59	2.75	2.59	2.62	2.71
4.0	3.44	3.12	3.13	3.34	3.58	3.32	3.36	3.50
5.0	4.20	3.75	3.74	4.07	4.38	4.01	4.05	4.26
6.0	4.93	4.35	4.28	4.78	5.16	4.67	4.69	5.01

[a] A, lognormal 30% error; B, lognormal 40% error; C, contaminated lognormal 40% error; and D, normal 40% error.

previous section. The distributions of estimated doses (tissue kerma) for each city are given in Table II. An important aspect of these distributions is that the relative numbers of survivors decrease more rapidly with increasing dose in Hiroshima than in Nagasaki.

The cohort contains a large number of persons at essentially zero dose, used as a comparison group. The immediate concern here is only with those having positive doses, and calculations below are based on the distribution of z among those above 0.10 Gy. The bias in risk estimates due to dose-estimation errors does depend on the size of the comparison group. This is a general phenomenon in regression with errors in covariables; information about the value of the intercept decreases the bias in the slope estimate. In the approach here this is accounted for automatically.

Recall that a parametric form is to be chosen for $f(x)$, namely that the probability of the true dose exceeding any value x is given by $\exp(-\theta_1 x^{\theta_2})$. For each city, estimates of the parameters θ_1 and θ_2 can be found such that the induced distribution of z matches very closely the observed distribution. The results depend only slightly on the error model used. For simplicity the θ-values resulting from a fifth "central" model, lognormal with 35% error, will be used for all analysis here. The results are:

$$\text{Hiroshima:} \quad \theta_1 = 2.84 \quad \theta_2 = 0.50$$

$$\text{Nagasaki:} \quad \theta_1 = 2.33 \quad \theta_2 = 0.50.$$

The values of θ_2 for the two cities are estimated, rather than constrained, to be equal. The fitted values are in approximate agreement to two decimal points.

With this city-specific model for $f(x)$ it is straightforward to compute Avg($x|z$) numerically from Eq. (1). Results for each error model at selected z-values are given in Table III.

The results shown in Table III will be discussed later in the paper, but it is briefly noted now that model C gives results similar to those for model B. The error models C and D are not pursued further, and most subsequent numerical results will be given only for the lognormal models, with the addition of an intermediate choice with 35% error to be discussed later. Note that the adjustments are greater for Hiroshima than Nagasaki; this is because the numbers of survivors decrease more rapidly with increasing dose in Hiroshima.

Convenient formulas for Avg($x|z$) can be obtained by fitting second-degree polynomials in log(z) to the *reduction factors*, $[z - \text{Avg}(x|z)]/z$. These formulas agree with the numerical computations of Avg($x|z$) to within 0.01 Gy over the range to 6 Gy. Coefficients for these approximations to the reduction factors are given in Table IV. It is suggested that no adjustments be made to estimates under 0.5 Gy, since they would be very small, and that estimates above 6 Gy be reduced to that value before adjustment. This truncation at 6 Gy has been made in all previous analyses of these data, since the small chance of survival suggests that estimates above 6 Gy are likely to be too large. It could

TABLE IV
Coefficients for Calculating Reduction Factors $[z - \text{Avg}(x|z)]/z$ as Polynomials in log(z)

	Error model lognormal percentage error	Coefficients for terms: Constant	log(z)	[log(z)]^2
Hiroshima	30%	0.03597	0.05807	0.01166
	35%	0.04732	0.07623	0.01336
	40%	0.06036	0.09684	0.01314
Nagasaki	30%	0.01500	0.05304	0.00885
	35%	0.01900	0.06545	0.01374
	40%	0.02817	0.08031	0.01558

TABLE V
Squared Coefficient of Variation for the Four Error Models

Estimated kerma (Gy)	$\text{Var}(x\mid z)/[\text{Avg}(x\mid z)]^2$							
	Hiroshima				Nagasaki			
	A[a]	B	C	D	A	B	C	D
0.5	0.090	0.158	0.137	0.181	0.091	0.160	0.140	0.194
1.0	0.088	0.154	0.133	0.154	0.089	0.157	0.135	0.170
2.0	0.086	0.148	0.135	0.124	0.087	0.152	0.132	0.141
3.0	0.085	0.144	0.142	0.108	0.086	0.148	0.135	0.124
4.0	0.084	0.141	0.152	0.097	0.085	0.145	0.139	0.112
5.0	0.083	0.139	0.164	0.089	0.084	0.143	0.145	0.104
6.0	0.082	0.137	0.177	0.083	0.083	0.141	0.152	0.097

[a] A, lognormal 30% error; B, lognormal 40% error; C, contaminated lognormal 40% error; and D, normal 40% error.

be argued that with proper adjustment for dose-estimation errors, this arbitrary procedure should be discontinued. The recommendation above is made so that the error model used will not have to apply to extremely high estimates. The DS86 estimates of tissue kerma are lower than previous estimates, which results in fewer survivors (presently 76) with estimates above 6 Gy tissue kerma.

For calculation of adjusted organ doses it is recommended, largely for simplicity, that organ dose estimates be adjusted by these reduction factors computed from tissue kerma estimates.

In the previous section several reasons were noted for interest in $\text{Var}(x\mid z)$, the variance of x among those at given z; and in $\text{Avg}(x^2\mid z)$, which is given by $\text{Var}(x\mid z) + [\text{Avg}(x\mid z)]^2$. Values of the squared coefficient of variation, $\text{Var}(x\mid z)/[\text{Avg}(x\mid z)]^2$, are given in Table V for the four error models.

Although the values in Table V can also be fitted very well by convenient approximations, they are remarkably constant in z for each error model and it would be an overinterpretation of the models used here to emphasize the small variation seen. The form of the additional variation in the chromosome aberration data due to errors in dose estimates depends on these values, as indicated in Eq. (5). Further, adding 1.0 to the numbers in Table V gives the ratio of $\text{Avg}(x^2\mid z)$ to $[\text{Avg}(x\mid z)]^2$, providing the quadratic covariable to be used for fitting linear-quadratic models. Since these ratios are nearly constant in z, simply regressing on $\text{Avg}(x\mid z)$ and $[\text{Avg}(x\mid z)]^2$ provides essentially the correct fit for a linear-quadratic model. That is, the fitted curve and significance tests for nonlinearity will be correct, but the coefficient of $[\text{Avg}(x\mid z)]^2$ will differ from the one appropriate for $\text{Avg}(x^2\mid z)$ by the reciprocal of the ratio discussed above in relation to the quadratic covariable.

The standard deviations of x given z are of some direct interest in understanding how informative the estimated doses are. However, the standard deviations do not give a clear description of variability, since the distributions $f(x\mid z)$ are far from normal. Thus graphs of some examples of the distributions $f(x\mid z)$ are given in Fig. 1. Also given in the same figures are corresponding graphs of the distributions $f(z\mid x)$. Note that the horizontal axis for each figure corresponds to x for one graph, and z for the other. The pairs of graphs are given together to emphasize the important distinction between the distribution $f(x\mid z)$ of true dose among those with a given estimated dose and the distribution $f(z\mid x)$ of estimated doses among those at the same true dose. These are the more realistic versions of the distributions indicated in the rows and columns of Table I.

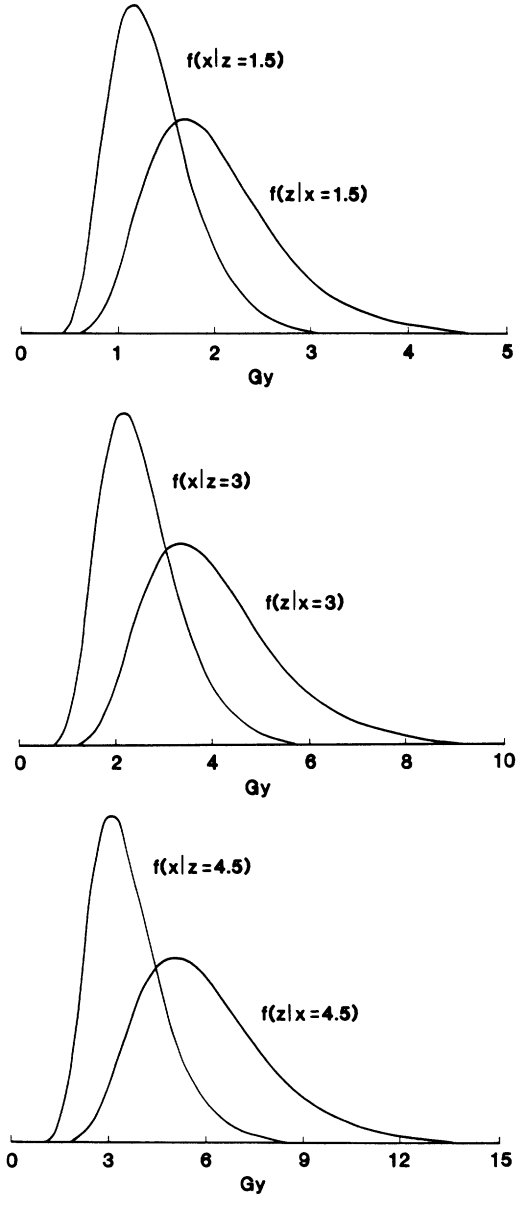

FIG. 1. Graphs of $f(z\mid x)$, and $f(x\mid z)$ for Hiroshima, for the lognormal 35% errors model.

TABLE VI
Increases in Linear Cancer Risk Estimates for Three Lognormal Error Models

	Lognormal percentage error	Increase in linear risk estimate (%)	
	Range of analysis:	0–6 Gy	0–4 Gy
All cancer except leukemia	30	10.0	6.8
	35	13.3	9.0
	40	16.7	11.4
Leukemia	30	6.1	4.3
	35	8.1	5.6
	40	10.2	7.2

Returning to consideration of Table III, the extent to which values of $\text{Avg}(x \mid z)$ are less than z fails to clarify how much bias in risk estimates will be removed by the methods of this paper, since the relative importance of various z-values in fitting dose–response models is not easily described. For the same reason, the sensitivity of results to the choice of error model is not made clear. Thus it is necessary to see how the methods of this paper affect final results, for various error models.

The LSS cancer mortality data have been reanalyzed using the data and models recently presented by Preston and Pierce (*15*). Table VI shows the resulting increase in their linear risk estimates, for all cancers except leukemia and for leukemia. These are in terms of relative risk for nonleukemia and absolute excess risk for leukemia. The risks depend on sex and age-at-exposure, and the summary measure of their paper is used, averaging over sex and three categories of age-at-exposure. Estimates are given for both the 0–4 Gy and the 0–6 Gy dose (tissue kerma) ranges. Preston and Pierce (*15*) gave both of these because of an apparent plateau in excess risk above 4 Gy. Here we have the additional motivation of seeing how restricting the dose range affects biases due to errors in dose estimates.

For the cancer data the standard errors of the parameter estimates relative to the estimates, i.e., the coefficients of variation, are essentially unchanged by the analysis allowing for dose-estimation errors. This is not happenstance, but reflects the fact that the additional variation in the data due to dose-estimation errors, beyond ordinary sampling variation, is negligible. For the chromosomal data, on the other hand, this is not the case and analysis allowing for dose estimation errors will lead to larger standard errors than those resulting from a binomial model. However, as discussed earlier, methods allowing for overdispersion relative to the binomial model have already been in use and the standard errors under the proposed method will be similar to those given by current methods.

DISCUSSION

As noted at the outset, the method developed here is not fundamentally different from that of Gilbert (*3*), but the implementation is more suitable for many purposes, especially since it is not linked to extensive calculations of the average true doses corresponding to a particular choice of estimated dose categories. The principles are also consistent with those discussed by Prentice (*8, 9*), who deals more specifically with application to survival data, but we have emphasized more some details specific to the LSS setting. The general approach laid out by Clayton (*11*) seeks to establish a unified solution to the problem of inference in the three components of the model: the dose–response model, the errors in doses, and the distribution of doses. This may have some theoretical advantages, but the methods involved need more development.

Much of the statistical literature on errors in covariables emphasizes reduction of bias by adjusting the "naive" estimates, i.e., those obtained by regression analysis using the observed covariables. The very different approach taken here has a number of important advantages for the needs associated with the RERF data. It lends itself better to the ongoing analyses of different end points and on different subcohorts of the data. The approach makes particularly clear the role of the distribution $f(x)$ of true doses, and how the analysis here should be modified to account for additional covariables and aspects such as shielding category. In regard to analysis of subcohorts, it becomes more clear that it is the distribution of doses $f(x)$ in the entire cohort, rather than in the subcohort under analysis, which is relevant to the correction procedure. Further, explicit consideration of the nonlinear nature of $\text{Avg}(x \mid z)$ is critical in analyses of the shape of the dose–response curve; see Pierce and Vaeth (*1*).

Further research on issues involving estimation of the distribution $f(x)$ of true doses may be useful. The final statistical methods suggested here make no allowance for errors in this estimation. The choices of $f(x)$ used here induce distributions of dose estimates which closely agree with the observed distributions, but the extent to which other choices of $f(x)$ would also do this has not been investigated.

The major difficulty, however, involves the assumptions regarding the statistical model $f(z \mid x)$ for errors in dose estimates. The results in Tables III and V indicate that there is very little difference between the use of the lognormal model and the contaminated lognormal model, which has substantially "heavier" tails. Thus, although the latter model may have substantial appeal as being more realistic, there seems to be little need to introduce the additional complexity. The normal model with a constant coefficient of variation does give somewhat different results, in the sense of smaller adjustments. We feel, as do others who

have considered this problem, that the lognormal model is preferable both on grounds that are specific to the dosimetry system, and on more general statistical considerations.

Important improvements may result from better assessment of the magnitude of dose estimation errors, and thus further study of this aspect of the DS86 system by RERF and dosimetry experts should be given high priority. This study should include investigation of the dependence of errors on shielding category, and distinction between random and systematic errors. Care must be taken to distinguish between errors as described by $f(z|x)$ and those described by $f(x|z)$.

The results in Table VI indicate not only the basic extent of sensitivity to the error model, but the degree to which the effect of errors in dose estimates is smaller when those with doses above 4 Gy are eliminated from the analysis. However, as discussed previously (*1, 15*), effects of errors in dose estimates are not the only motivation for this restriction. Other critical issues involved are that the true dose-response may not be linear over a wide range, and that risk estimation at low doses is the predominant concern. Restrictions of the dose range have also been used to lessen the effects of dosimetry errors. The methods of this paper are a more direct approach to this, but such restrictions will still help by making the choice of a dose error model less critical.

It should be noted that there is another important way of looking at the entire problem addressed in this paper. The dosimetry system takes no account of the information provided by the fact that the individuals *survived.* That is, the estimates depend only on survivor location and shielding, along with physical calculations based on yield of the bombs and radiation transport. If a model for the probability of surviving as a function of true dose were available, it would be possible to make dose estimates based on the survivor location and shielding, *and* the information provided by survival. It is not suggested that this should actually be done. The way that estimated distributions $f(x)$ of true doses are used in this paper essentially incorporates the information provided by survival, in a more feasible way. This point is raised primarily to reinforce the statistical reasoning that, except for those at very small doses, the likelihood that true dose is less than estimated dose is greater than the likelihood that true dose is larger than estimated dose.

There are issues in addition to bias in dose–response analyses which are related to errors in dose estimates. For example, there is interest in the extent of biological variation among individuals in their general sensitivity to irradiation. This is relevant to the estimation of cancer risks, since substantial variation of this type might imply that those in the study cohort, having survived the acute effects, would be less sensitive to radiation-induced cancer than a general population. This has been investigated by looking for positive associations between different radiation effects (cancer, acute symptoms, chromosomal aberrations, etc.) among those at the same dose level; see, for example, Neriishi *et al.* (*16*). It is certain, though, that random errors in dose estimates will induce spurious positive association of this nature. Among those at (approximately) the same estimated dose, there is in fact a distribution of true doses, and two types of radiation-induced effects will tend to occur together for those with higher true doses. Thus investigations must consider whether positive associations found are greater than would be due to this. The statistical formulation of this report can be extended such calculations.

In conclusion, although further research in this area is desirable, it is important to begin using the methodology developed here. In some analyses of the RERF data it will be important to investigate the sensitivity of conclusions to assumptions about the error model. This would be particularly important, for example, in analyses of the shape of the dose response and in studies of individual variation in sensitivity to irradiation. Many reports, however, are already complicated by the necessary attention to a large number of issues, and such special consideration of error models will be impractical. These investigations ordinarily use linear dose response models, and for that case the results here suggest that the uncertainty in risk estimates due to choice of an error model may be relatively small in relation to other uncertainties. For such analyses it may be best to rely primarily on the use of a given choice of an error model, with some brief indication of the effect of the adjustment.

The authors thus feel that it is useful to suggest an error model to be used at this time, for analyses where it is infeasible to devote substantial special attention to the issue. That is, it will be better to make adjustments with some error model than none at all. It seems clear that lognormal error models should be used for this purpose, and it seems to the authors that the choice of the 30% error model may be somewhat optimistic. A primary basis for this is the judgment of Jablon (*2*) that "sources identified in this section amount to about 30% plus or minus," preceded by the comment "reflecting on how they [certain aspects] were derived it seems not unlikely that we have underestimated." The uncertainties that Jablon considered were largely those due to inaccuracies in the assessment of the survivors' location

TABLE VII
Representative Values of Adjusted Dose Estimates, Lognormal 35% Error Model

Estimated kerma (Gy)	0.50	1.00	2.00	3.00	4.00	5.00	6.00
Adjusted kerma (Gy)							
Hiroshima	0.50	0.95	1.79	2.56	3.28	3.98	4.64
Nagasaki	0.50	0.98	1.86	2.68	3.46	4.20	4.92

and shielding, inputs which are common to the old and the new dosimetry systems. Errors in DS86 somewhat smaller than 30% have been discussed (Ref. *12,* Chapter 9), but that assessment apparently takes the input parameters of survivor location and shielding at face value.

The authors suggest that for the time being some special focus be placed on the lognormal 35% error model. Numerical values for dose adjustments under this model were not given in Table III, although the formula is given in Table IV, so some representative values are given in Table VII. The values corresponding to Table V, for this model, are in the range 0.11–0.12.

ACKNOWLEDGMENTS

The RERF is supported equally by the Japanese Ministry of Health and Welfare, and the U. S. National Academy of Sciences through a contract with the Department of Energy. The first author's work was done while on leave from Oregon State University, and his research on the statistical methods was also supported by a grant from the National Science Foundation. The authors acknowledge the major contribution of anonymous reviewers for both the RERF Technical Report version and this revision, and also valuable assistance with the statistical theory from Daniel Schafer, Oregon State University.

RECEIVED: April 18, 1989; ACCEPTED: April 30, 1990

REFERENCES

1. D. A. PIERCE and M. VAETH, *The Shape of the Cancer Mortality Dose–Response Curve for Survivor Data.* Technical Report 7–89. Radiation Effects Research Foundation, Hiroshima, 1989.
2. S. JABLON, *Atomic Bomb Radiation Dose Estimate at ABCC.* Technical Report 23-71. Atomic Bomb Casualty Commission, Hiroshima, 1971.
3. E. S. GILBERT, Some effects of random dose measurement errors on the analyses of atomic bomb survivor data. *Radiat. Res.* **98,** 591–605 (1984). [RERF TR 12-82]
4. W. G. COCHRAN, Errors of measurement in statistics. *Technometrics* **10,** 637–666 (1968).
5. M. H. GAIL and D. P. BYAR (Eds.), Workshop on Errors-in-Variables: NIH. *Stat. Med.* **8**(9) (1989).
6. W. A. FULLER, *Measurement Error Models.* Wiley, New York, 1987.
7. B. ARMSTRONG, Measurement error in the generalized linear model. *Commun. Statist.* **14,** 529–544 (1985).
8. R. L. PRENTICE, Covariate measurement error and parameter estimation in Cox's regression model. *Biometrika* **69,** 331–342 (1982).
9. R. L. PRENTICE, Covariate measurement errors in the analysis of cohort and case-control studies. In *Survival Analysis* (J. Crowley and R. Johnson, Eds.), Vol. 2. Institute of Mathematical Statistics, 1982.
10. R. L. PRENTICE, Binary regression using an extended β-binomial distribution, with discussion of correlation induced by covariate measurement errors. *J. Am. Statist. Assoc.* **81,** 321–327 (1986).
11. D. G. CLAYTON, Models for the analysis of cohort and case-control studies with inaccurately measured exposures. In *Statistical Models for Longitudinal Studies of Health* (J. H. Dwyer, Ed.). Oxford University Press, Oxford, 1988.
12. W. C. ROESCH (Ed.), *Final Report of US–Japan Joint Reassessment of Atomic Bomb Radiation Dosimetry in Hiroshima and Nagasaki.* Radiation Effects Research Foundation, Hiroshima, 1987.
13. D. A. PIERCE, D. O. STRAM, and M. VAETH, *Allowing for Random Errors in Radiation Exposure Estimates for the Atomic Bomb Survivor Data.* Technical Report 2-89, Radiation Effects Research Foundation, Hiroshima, 1989.
14. D. L. PRESTON, M. E. MCCONNEY, A. A. AWA, K. OHTAKI, M. ITOH, and T. HONDA, *Comparison of Dose–Response Relationships for Chromosome Aberration Frequencies between the T65D and DS86 Dosimetries.* Technical Report 7-88. Radiation Effects Research Foundation, Hiroshima, 1988.
15. D. L. PRESTON and D. A. PIERCE, The effect of changes in dosimetry on cancer mortality risks estimates in the atomic bomb survivors. *Radiat. Res.* **114,** 437–466 (1988).
16. K. NERIISHI, D. O. STRAM, M. VAETH, S. MIZUNO, and S. AKIBA, *The Observed Relationship between the Occurrence of Acute Radiation Sickness Symptoms and Subsequent Cancer Mortality in the Hiroshima and Nagasaki Data.* Technical Report 18-89. Radiation Effects Research Foundation, Hiroshima, 1989.