A Method for Estimating Occupational Radiation Dose to Individuals, Using Weekly Dosimetry Data

Toby J. Mitchell*,† George Ostrouchov*,‡ Edward L. Frome*, and George D. Kerr**

Mitchell, T. J., Ostrouchov, G., Frome, E. L., and Kerr, G. D. A Method for Estimating Occupational Radiation Dose to Individuals, Using Weekly Dosimetry Data. *Radiat. Res.* **147**, 195-207 (1997).

Abstract: Statistical analyses of data from epidemiologic studies of workers exposed to radiation have been based on recorded annual radiation doses. It is usually assumed that the annual dose values are known exactly, although it is generally recognized that the data contain uncertainty due to measurement error and bias. We propose the use of a probability distribution to describe an individual's dose during a specific period of time and develop statistical methods for estimating this distribution. The methods take into account the "measurement error" that is produced by the dosimetry system, and the bias that was introduced by policies of recording doses below a threshold as zero. The method is applied to a sample of dose histories over the period 1945 to 1955 obtained from hard copy dosimetry records at Oak Ridge National Laboratory (ORNL). The result of this evaluation raises serious questions about the validity of the historical personnel dosimetry data that is currently being used in studies of the effects of low doses in nuclear industry workers. In particular, it appears that there was a systematic underestimation of doses for ORNL workers. This may result in biased estimates of dose-response coefficients and their standard errors.

Keywords: Bayesian estimation, likelihood, film dosimeter, missing dose, dose uncertainty

INTRODUCTION

Pocket ionization chambers (or pocket meters) were initially considered the primary device for monitoring personnel exposures, with a film dosimeter being only a valuable adjunct (1). With expanding experience at Oak Ridge, Tennessee, and with the startup of the production facilities at Hanford, Washington, in 1944, this practice was reversed, and the film dosimeter provided the official dose of record, while the pocket meter became the day-to-day means of monitoring personnel exposures in the workplace (2). At ORNL, however, the daily pocket-meter readings were also maintained as a part of an individual's dose records (1).

An individual's radiation dose of record at ORNL for external penetrating radiation, principally gamma rays, is based on pocket meters from 1943 to July 1944, film badges from then to 1975, and thermoluminscent dosimeters since 1975 (3). The pocket meters were evaluated daily (minimum detectable limit of 0.02 mSv), and the film badges were evaluated weekly from July 1944 to July 1956, when quarterly monitoring was initiated (minimum detectable limit of 0.30 mSv). This is the period to which the methods developed in this report are applied. Several reports have already been published about missing dose during the weekly evaluations of film badges (4, 5, 6, 7). However, there is considerable doubt in the current literature concerning the lower detection limit of the film-badge dosimeters (8). The general issue of uncertainty in individual dose estimates in epidemiologic studies of nuclear industry workers has also been discussed (9, 10, 11, 12).

The lower limit of detection of the most sensitive film used at ORNL was 0.10 to .30 mSv. A lower detection limit of 0.10 mSv was possible if an experienced technician evaluated the exposed films with special care (13). During film-badge exchange, when hundreds to thousands of films were read in large batches by technicians with widely varying experiences, a lower limit of detection of about 0.30 mSv was about as good as could be expected (14). In practice, a film-badge reading of zero means the radiation dose to the worker was less than 0.30 mSv unless a smaller value is given. Thus, the missing dose from weekly evaluations of the film badges occurred primarily among those workers

*Mathematical Sciences Section, Computer Science and Mathematics Division, Oak Ridge National Laboratory, P. O. Box 2008, Oak Ridge, Tennessee 37831-6367

†Deceased

‡Author to whom correspondence should be addressed.

**Assessment Technology Section, Health Sciences Research Division, Oak Ridge National Laboratory, P. O. Box 2008, Oak Ridge, Tennessee 37831-6383

with the lowest radiation-dose estimates between 1944 and 1956 (4, 5, 6).

The design of the film badge and its use at ORNL changed considerably over the years. In November 1951, for example, the photo film badge was introduced and all ORNL employees were required to wear a film badge on the job (1). Prior to November 1951, only those ORNL employees who entered a radiation area were required to wear a film badge. Two or more filters were used in all ORNL film badges to aid in interpreting the radiation dose and in resolving the difficulty due to the fact that the unshielded films were more sensitive to x rays between 50 and 100 keV than to x or gamma rays above 200 kev (14). The film-badge readings quoted throughout this report are estimates of the equivalent dose from external penetrating radiation at a depth of approximately 1 cm within the total body or a major portion of the total body.

Typically, epidemiological studies of the effects of external penetrating radiation on worker health have relied on recorded annual doses to the individuals in the population. These annual doses were obtained by adding up recorded weekly film-badge readings. In the statistical analysis, these annual dose values have been treated as though they are known exactly, although everyone recognizes that there is uncertainty due to measurement error and bias. It is usually assumed that the measurement errors "average out" and that the bias is small.

This project is motivated by the need for adjustment for dose bias and uncertainty in epidemiologic dose-response analyses. The first step is an adjustment for bias and quantification of uncertainty in dose estimates, which is the subject of this paper. Except for some general remarks, the effect of dose bias and uncertainty on dose-response estimates is not considered (15).

The objective of this report is to provide methodology for estimating the true dose of an individual during a year, given the recorded weekly exposure histories for that individual in that year. The "true" dose of an individual is considered to be the quantity of radiation encountered by the film badge(s) worn by the individual during the period in question. The relevant and difficult issues involved in estimating the dose to the body (or, even more relevant and more difficult, the estimation of doses to specific organs) will not be considered here but can be based on the same framework.

The dose estimate proposed for each individual is a nonparametric probability distribution. This is the most general description of uncertainty and can be reduced to other descriptions of uncertainty. A nonparametric probability distribution estimate, consisting of many (say 100) density points, can be reduced to a more concise description such as the five points of a boxplot (see Section on Dose Estimates), or to a few parameters of an assumed distribution (such as the mean and variance of a normal or a lognormal distribution). Each reduction is a loss of some information and a gain in simplicity. These can be computed for an individual or for any cohort of individuals. Such generality allows the dose estimates to be useful for many purposes including adjustment for dose uncertainty in epidemiologic dose-response analyses by methods already known or yet to be developed.

A more detailed version of this report and related computational methods are available (16, 17) (or URL http://www.epm.ornl.gov/~ost/compstat.html).

METHOD FOR A SINGLE PERIOD

A Bayesian statistical approach is used to estimate the unobserved quantities (true doses) given the values of the observed ones (recorded doses). A relationship between the true dose and the recorded dose in the form of a conditional probability distribution is the key element of the method.

For each dose, the estimate is expressed in the form of a probability distribution. A point estimate (single "best" value, by some criterion) could be obtained from this distribution, but we shall avoid this, since we regard the probability distribution itself as the estimate, and think of any reduction as a loss of information. In particular, if annual doses are to be used as inputs to a model that relates health effects to radiation dose, it is necessary to obtain point estimates and to quantify the uncertainty in these values.

In Bayesian estimation, quantities of interest, observed and unobserved, are endowed with a joint prior probability distribution that represents (approximately) the state of knowledge about them prior to (or external to) observation or measurement. Then the actual values of the observed measurements are put in, as conditioning information, and the laws of probability are used to find the conditional distribution of the unobserved values *given* the observed ones. See for example (18) for further background on Bayesian estimation or (19, 20) for an application in dosimetry.

In a single exposure period (e.g., one week), there are two quantities of interest:

x the unobserved true dose to the film badge, and

z the recorded dose to the film badge.

The "functional" approach to measurement errors is used because we consider the unobserved x to have a fixed value (21). Nevertheless, x is treated as a random variable to express the uncertainty associated with our knowledge of its true fixed value. For example, there can be only one true value for x, but, in the absence of knowledge of what that value is, we attach a (prior) probability $P(x)$ to every possible value of x, where $\sum_x P(x) = 1$. We shall refer to the function $P(x)$ as the *probability distribution* of the random variable x. The interpretation of probability here is degree of belief in the truth of the proposition that the true dose is x. This interpretation provides a mathematical representation of the degree of uncertainty about deterministic quantities: a small bit of probability placed at each of a large number of values of x reflects a high degree of uncertainty, whereas a probability of 1 placed at a single value reflects complete certainty.

We emphasize that $P(x)$ refers to the distribution of probabilities that concerns **one individual in one exposure period**. This is important to note, because in other literature distributions often refer to a cohort of individuals.

The recorded dose z is also treated as a random variable. Prior to its observation, for a known x, there is uncertainty in its value. This allows the assumed relationship between z and x to take the form of a conditional probability distribution $P(z|x)$. This is an "if x, then z" relationship, but with uncertainty built in, uncertainty that exists *prior* to the observation of z.

The language of probability is used to arrive at a statement about x *given* z. The conditional probability distribution $P(x|z)$ is called the *posterior* distribution and is given by the Bayes' Theorem (see (18), for example)

$$P(x|z) = c(z)P(x)P(z|x), \tag{1}$$

where $c(z)$ is a normalizing constant which ensures that $\sum_x P(x|z) = 1$.

The key component for implementing this approach is $P(z|x)$. In effect, $P(z|x)$ is the answer to the question: "If the true dose is x, what is the probability that the recorded value is z?" This is determined by careful consideration of the properties of the measuring device (in this case the film badge and the system used in reading and recording its dose). Note that that this is a function of two variables, namely x and z, and it is constructed by specifying a distribution on z for each possible (fixed) value of x (as we do in the section that follows). After specifying $P(z|x)$ for all possible z and x, it is used as a function of x for each observed z. This is the "likelihood" of x for the observed z and is denoted by $L(x|z)$.

The prior distribution $P(x)$ is less critical but more problematic. What beliefs and uncertainties should go into the determination of the prior probabilities? In most situations, however, it is possible to formulate a description of $P(x)$ that is acceptably objective.

Constructing the Likelihood Function for the 1945–1955 ORNL Cohort

The likelihood $L(x|z)$ is available from the complete specification of $P(z|x)$ for all possible values of x and z. We begin by constructing $P(z|x)$.

Let $\tilde{z}$ be the *expressed dose* to the badge; that is, the reading that would be recorded if there were no rounding or censoring. ("Rounding" means that readings are given in multiples of 0.05 mSv. "Censoring" is the practice of recording as zero all readings that are below a certain threshold.) The variability in $\tilde{z}$ for fixed x is intended to represent instrument error and reading error. We assume that $\tilde{z}$ has a lognormal distribution such that $\log(\tilde{z})$ has mean $\log(x)$ and standard deviation $\alpha(x)$, **both of which depend on** x. Thus,

$$P(\tilde{z}|x) = \frac{1}{\sqrt{2\pi}\tilde{z}\alpha(x)} \exp\left\{\frac{1}{2\alpha(x)^2}\left[\log(\tilde{z}) - \log(x)\right]^2\right\}.$$

Information about the dependence of α on x is obtained from (13):

> "... in ordinary routine procedures using technicians to process the film badges, the probable error is about ± 30 mrad [or 0.3 mSv of γ-dose] ... This ± 30 mrad does not represent our total probable error in reading the film badges, except when the readings are from 0 to 30 mrad. If the exposure is to hard γ-radiation [.1 to 3 MeV], we can read 100 mrad to ± 0.015 rad [i.e., $\pm 15\%$] or 1 rad to ± 0.1 rad [i.e., $\pm 10\%$]."

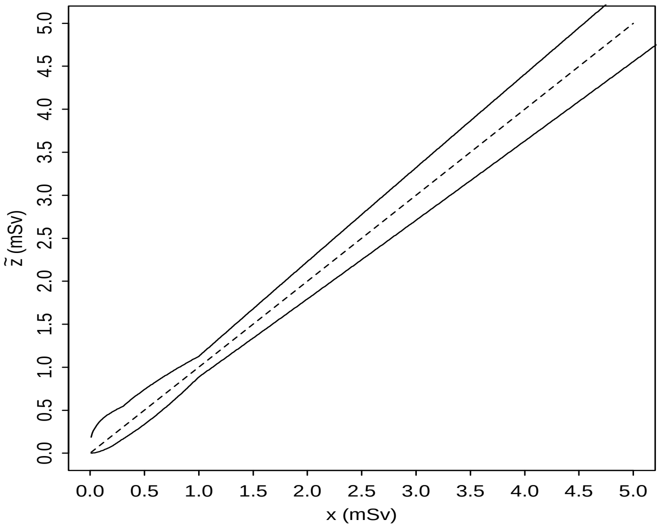

Figure 1: 99% probability bounds on expressed dose ($\tilde{z}$) versus true dose (x).

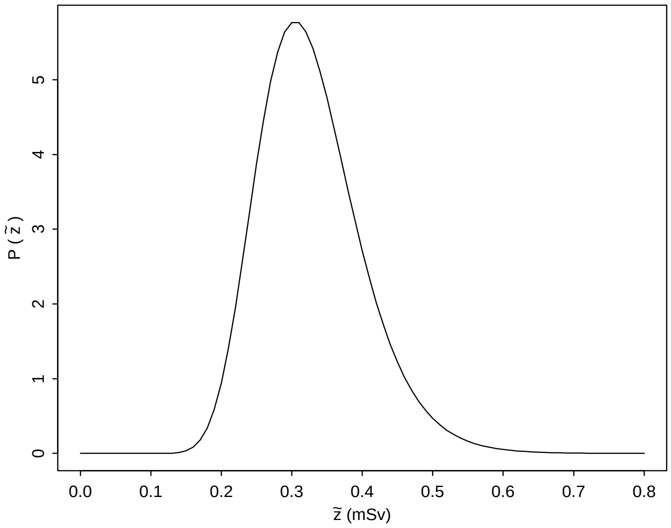

Figure 2: Probability density function of expressed dose $\tilde{z}$ when true dose x is 0.32 mSv.

To translate this into reasonable values for $\alpha(x)$, it is assumed that the "upper 3-standard deviation" limit on $\log(\tilde{z})$ corresponds to the following upper limits on $\tilde{z}$: 0.3 mSv at $x = 0.01$ mSv, 0.6 mSv at $x = 0.3$ mSv, 1.15 mSv at $x = 1$ mSv and 11 mSv at x=10 mSv. Under the lognormal assumption, the probability that these upper limits are exceeded is only 0.0013. That is, we are treating Morgan's (13) "probable errors" essentially as maximum errors. This interpretation is consistent with Morgan's usage of them to compute the errors for sums of film-badge readings. By setting these upper limits to $xe^{3\alpha}$, i.e., the logarithms of the limits to $\log(x) + 3\alpha$, we find $\alpha = 1.134$ at $x = 0.01$ mSv, $\alpha = 0.231$ at $x = 0.3$ mSv, $\alpha = 0.0466$ at $x = 1$ mSv, and $\alpha = 0.0318$ at $x = 10$ mSv. To interpolate between these values, the following piecewise linear function in α vs. $\log(x)$ is used:

$$\alpha(x) = \begin{cases} -0.0884874 - 0.2654017 * \log(x) & 0 < x \leq 0.30 \\ 0.04658731 - 0.1532109 * \log(x) & 0.30 \leq x \leq 1.00 \\ 0.04658731 - 0.00643505 * \log(x) & 1.00 \leq x \leq 10.00 \end{cases} \quad (2)$$

Because the lognormal distribution is skewed to the right, the "probable error" in the negative direction is less than that in the positive direction. For example, the 99% probability bounds for $\tilde{z}$ are (0.0005, 0.19) at $x = 0.01$, (.17, 0.54) at $x = 0.30$, (.89, 1.13) at $x = 1.00$, and (9.21, 10.85) at $x = 10.00$. The bounds based on this interpolation are shown in Fig. 1, for $0.01 \leq x \leq 5.00$ mSv. Fig. 2 shows, as an example, $P(\tilde{z}|x)$ when $x = 0.32$. This is a lognormal distribution such that $\log(\tilde{z})$ has a normal distribution with mean $\log(.32) = -1.139$ and standard deviation $\alpha(.32) = 0.221$ (from Eq. (2)).

The above definition of expressed dose is used in this report. However, other more complex definitions are possible. For example, it is believed that some workers in higher dose categories might have removed their badges to avoid exceeding dose limits. This can be accounted for by an expressed dose distribution that is a mixture of a lognormal and another distribution that is skewed to the left. The mixing probabilities would be set by the probability of removing a badge. Such information is not available, but this could be used to asses sensitivity of the results to badge removals.

Recall that $\tilde{z}$ is the *expressed* dose. We assume in this section that the *recorded* dose z is obtained from $\tilde{z}$ by rounding to the nearest[1] multiple of 0.05 mSv, and reporting this value if it is greater than or equal to 0.30 mSv. If $\tilde{z}$, after rounding, is less than 0.30 mSv, then zero is reported. We call this last policy the *censoring convention*. These are simplified versions of the rounding and censoring that were done when the historical ORNL data were recorded. For example, the historical data contains a small number of nonzero values that are less than 0.30 mSv, which clearly were not censored.

Fig. 3 is derived from Fig. 2 by applying both the rounding and censoring conventions. This is the as-

[1]We use unequal probability rounding suggested by historical data. Multiples of 0.10 favor multiples of 0.05 at the rate of 85% to 15%.

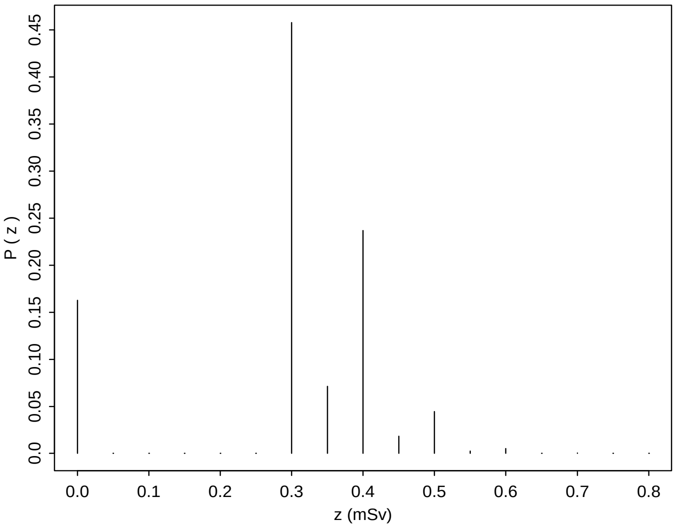

Figure 3: Probability density of recorded dose z when true dose x is 0.32 mSv.

sumed distribution of recorded doses that results if the true dose $x = 0.32$ mSv. It is interpreted as a description of the recorded-dose frequency distribution of a large number of film badges exposed to 0.32 mSv and processed through the dosimetry and recording system used at ORNL prior to 1956.

The above definition of $P(z|x)$ is used to compute the likelihood of any x for an observed z. For example, Fig. 4 shows the likelihood of x when $z = 0.40$ mSv, the recorded dose to the badge. This provides an answer to the question: "How likely is it that this observed value $z = 0.40$ mSv was a result of a possible true dose x?"

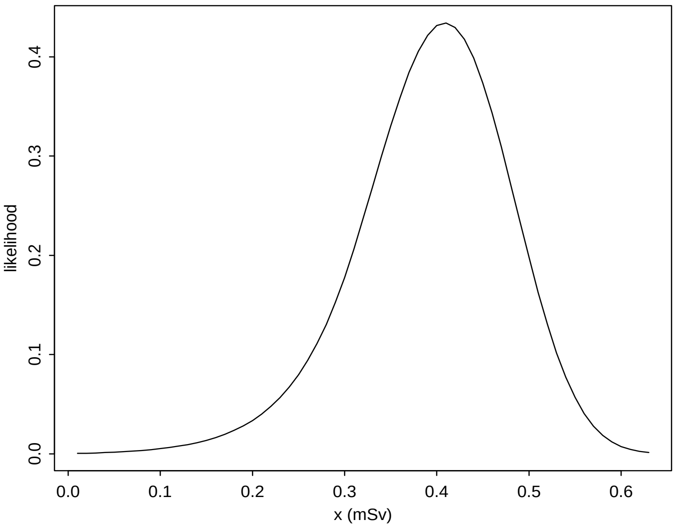

Figure 4: Likelihood of true dose x when recorded dose z is 0.40 mSv.

The Weekly Prior

The lognormal distribution is used as the prior distribution $P(x)$. The parameters μ and σ are the mean and the standard deviation of $\log(x)$. Note also that $\exp(\mu)$ is the median (or 50th percentile) of x. This distribution is chosen partly for computational convenience, and partly because it reflects the general belief that larger doses are less likely than smaller ones. This belief is consistent with the statistical distribution of film-badge readings observed in large homogeneous populations. Another good choice for the prior would be the gamma distribution, for example. The methodology can be applied with any choice of prior distributions (including improper distributions). The procedure used to specify the lognormal prior parameters is be discussed in a later section.

Computing an Individual Weekly Dose Estimate

Two examples of a weekly dose estimate are given: one from a recorded dose of 0.40 mSv and another from recorded dose of 0 mSv. For illustration purposes, a weekly prior with a median of 0.3 mSv and 75th percentile of 1 mSv is used. This prior is chosen because its median falls between the two recorded dose examples and it provides a good example of how data changes prior information. The center of the prior (the median) also falls on the censoring point. This gives lognormal parameter values of $\mu = -1.204$ and $\sigma = 1.821$. The 5th, 25th, 50th, 75th, and 95th percentiles of this distribution are 0.014, 0.084, 0.30, 1.0, and 6.0 mSv, respectively. Note that the spread of the distribution is quite wide, reflecting a considerable amount of prior uncertainty about the true dose x in a given week. This is called the "fixed" prior.

For computational convenience, the distributions and the likelihood are discretized so that all of the probability lies on multiples of a small dose (for example, 0.01 mSv). Thus, in our example, $P(x)$ is the probability that the lognormal variable with parameters $\mu = -1.204$ and $\sigma = 1.821$ lies within 0.005 mSv of x, for $x = 0.01, 0.02, 0.03, \ldots$. Details of the discretization, including spacing of the discrete mass points and their coverage, are discussed in (17). The reader who is interested in prior and likelihood construction and computation for another cohort may require the additional detail presented in (17).

Let $\mathbf{x}$ be the vector of discretized x values. Then $P(\mathbf{x})$ is the vector of associated prior probabilities and

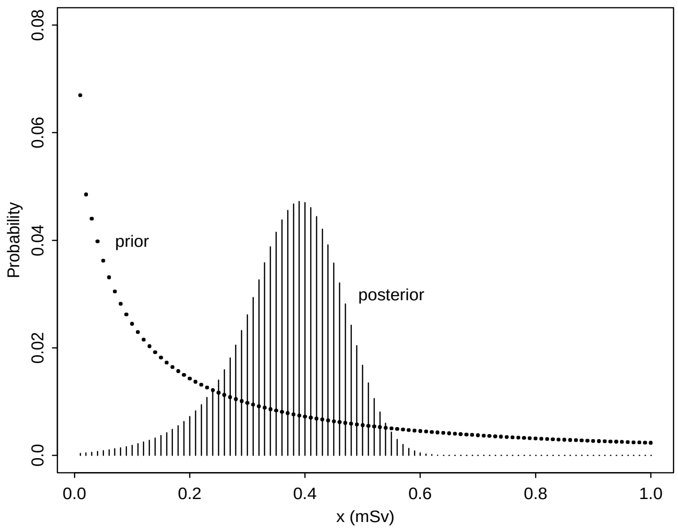

Figure 5: Comparison of prior to posterior distribution of true dose x when recorded dose z is 0.40 mSv.

$L(\mathbf{x}|z)$ is the vector of likelihood values. Equation (1) in terms of these vectors is

$$P(\mathbf{x}|z) = c(z)P(\mathbf{x})L(\mathbf{x}|z), \quad (3)$$

where the vector product is element-wise, meaning that the ith element of $P(\mathbf{x}|z)$ is $c(z)$ times the ith element of $P(\mathbf{x})$ times the ith element of $L(\mathbf{x}|z)$.

Suppose $z = 0.40$ mSv. The posterior distribution $P(\mathbf{x}|z = 0.40)$ is obtained by element-wise product of $P(\mathbf{x})$ and $L(\mathbf{x}|z = 0.40)$, and normalizing so that the sum of probabilities is one (See Eq. (3)). This is shown in Fig. 5 together with the fixed prior distribution $P(x)$ for comparison. The knowledge of a single film-badge reading $z = 0.40$ mSv considerably sharpens our knowledge of x in the sense of reducing the uncertainty about it.

Now suppose the recorded dose z had been zero rather than 0.40 mSv. The likelihood function in the case $z = 0$ is shown in Fig. 6. In Fig. 7, the fixed prior and the posterior distributions are plotted together for comparison. Note that although the film badge serves to exclude the possibility that the true dose is greater than about 0.40 mSv, it does not distinguish well among low values of x. In this low dose region, the posterior distribution essentially mirrors the prior distribution, and is sensitive to the particular choice of prior. Comparing Fig. 5 and Fig. 7, shows that the posterior distribution is much less sensitive to the prior when $z > 0$ than when $z = 0$.

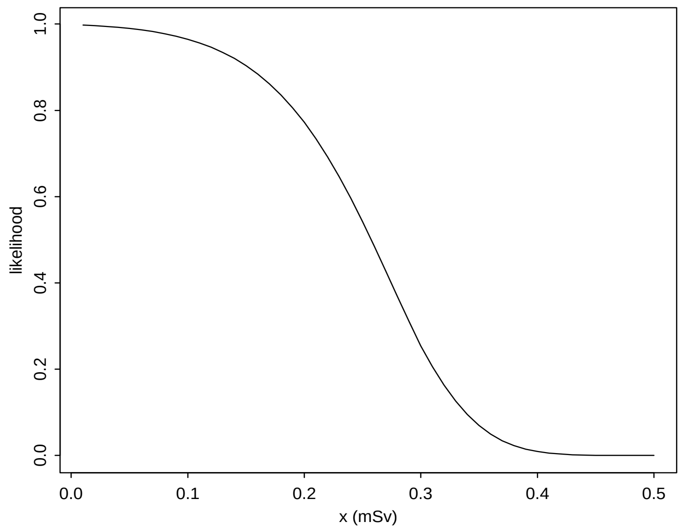

Figure 6: Likelihood of true dose x when recorded dose z is zero.

ESTIMATING CUMULATIVE DOSE

Now consider the estimation of a yearly dose based on a sequence of n film-badge readings, obtained weekly from a single individual. Let $x_T = \sum_{i=1}^{n} x_i$ where x_i is the true dose to the badge worn by the individual during the i^{th} week, and let z_i be the corresponding recorded dose. Also, let the set of n true doses be $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$, and the set of recorded doses be $\mathbf{z} = \{z_1, z_2, \ldots, z_n\}$. The objective is to obtain the posterior distribution $P(x_T|\mathbf{z})$, which will serve as an estimate of x_T. The simplest approach is to treat each

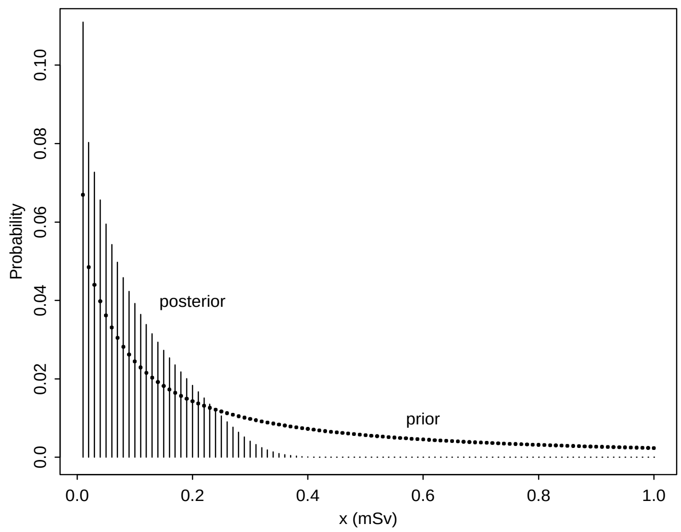

Figure 7: Comparison of prior to posterior distribution of true dose x when recorded dose z is zero.

week as independent of the others, so that

$$P(\mathbf{x}|\mathbf{z}) = \prod_{i=1}^{n} P(x_i|z_i). \qquad (4)$$

That is, conditional on the set of recorded doses $\mathbf{z}$, the set of true doses can be treated as independent random variables, whose individual distributions are given by $P(x_i|z_i)$. The distribution of the sum x_T can be obtained numerically in various ways (see (17)). For this report, a large random sample from this distribution is generated and a histogram or a boxplot of the sample is reported. To generate x_{Tm}, the mth value of x_T in this sample, first the set $\{x_{1m}, x_{2m}, \ldots, x_{nm}\}$ is generated by drawing x_{im} from $P(x_i|z_i)$ for $i = 1, \ldots, n$. These values are then summed:

$$x_{Tm} = \sum_{i=1}^{n} x_{im},$$

and the procedure is repeated, until a large sample (consisting of several thousand values) is generated. Combining intervals through simulation allows nonuniform intervals as well as the introduction of serial dependence structures between intervals. In the case of serial dependence, Eq. (4) does not hold, but the simulation approach to computing the x_T distribution is still valid. For example, a dependence structure can be introduced by providing a dependence relationship of a prior in a given week on the posterior of the previous week.

Note that $P(x_i|z_i)$ is represented by a discrete approximation on a finite number of points. The inverse cumulative distribution function method is used (see (22), for example) for drawing sample points from this discrete distribution. See (17) for more detail.

The result of this procedure, is demonstrated using weekly records from two individuals. The sequence of recorded film-badge readings ($\mathbf{z}$) for person A in 1948 is:

{0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.35, 0, 0, 0, 0.30, 0, 0, 0, 0, 0, 0, 0, 0, 0},

the total for the year being 0.65 mSv.

For each week, $P(x_i|z_i)$ is computed, in a manner similar to the one used to generate Fig. 5 and Fig. 7. Then 5,000 values of x_T are generated as discussed above. A histogram[2] of this sample, which can be viewed as an estimate of $P(x_T|\mathbf{z})$, is shown in Fig. 8. This suggests that the true 1948 dose of person A is roughly between 3.5 and 7 mSv.

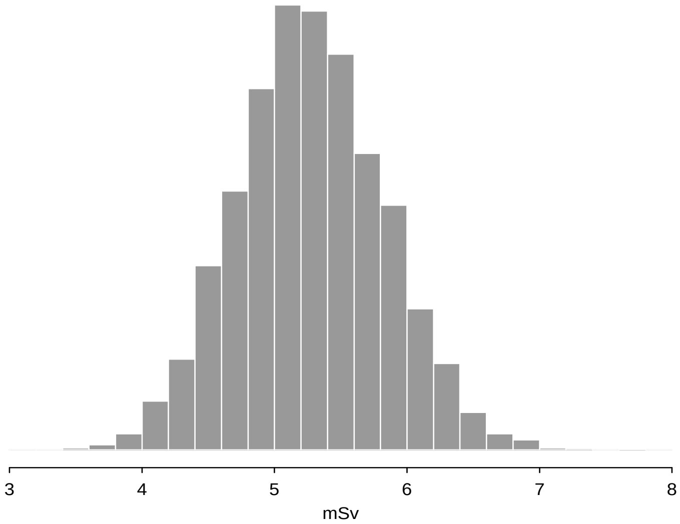

Figure 8: Yearly dose distribution estimate for person A (with recorded dose of 0.65 mSv) based on a fixed prior.

For person B, the sequence of recorded badge readings in 1954 is:

{ 2.1, 0, 0.8, 0.95, 1.4, 0.5, 0, 0.7, 0, 1.4, 0.3, 3.45, 1, 1.6, 0, 1.6, 0, 0.8, 0.5, 0.7, 0, 0.4, 0, 0, 0.6, 0.75, 0, 0, 0, 0, 0, 0, 0, 0, 1.5, 0, 0, 0, 0, 0, 0.9, 0.7, 1.1, 0, 0, 0.9, 0, 0, 0, 1.7, 0.55, 0.3 },

the total for the year being 27.2 mSv. Fig. 9 shows the estimate of $P(x_T|\mathbf{z})$ for this individual. It suggests that the true 1954 dose for person B is roughly between 27.5 and 31 mSv.

Note that for both individuals the estimate of yearly dose is centered higher than the sum of recorded doses. This is much more pronounced with the lower recorded annual dose.

USE OF ADDITIONAL DATA

Additional information can be incorporated into the construction of the likelihood and the prior. The available weekly data (described below in the application section) also includes a pocket-meter dose corresponding to each recorded film-badge reading. This weekly

[2] All histograms presented here are density estimates and as such have a total area of 1. The y-axis scale depends on the histogram bin size and is not displayed to avoid inappropriate comparisons.

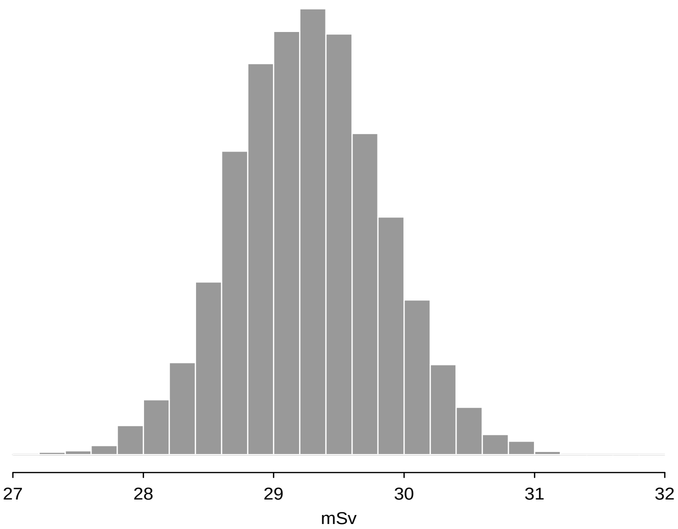

Figure 9: Yearly dose distribution estimates for person B (with recorded dose of 27.2 mSv) based on a fixed prior.

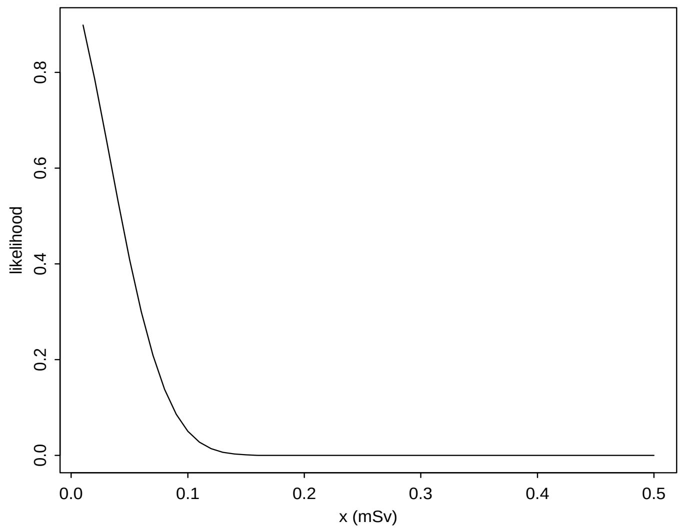

Figure 10: Likelihood of true dose x when recorded dose z is zero without censoring.

pocket-meter dose is calculated as the sum of minimum daily readings of a pair of pocket meters. This sum is the weekly pocket-meter dose and is denoted by w. Also, let $\mathbf{w} = \{w_1, w_2, \ldots, w_n\}$ be the vector of n weekly pocket-meter doses that corresponds to $\mathbf{z}$, the vector of recorded film-badge readings.

The weekly pocket-meter dose, w, is used to indicate whether censoring was needed for a given zero film-badge reading and to provide location information for a very diffuse prior. A more rigorous approach to the inclusion of pocket-meter data would require likelihood construction and prior construction for pocket meters by reviewing historical information on daily pocket-meter dose measurement at ORNL. These are the two basic components of our methodology that have to be "customized" in every new application. For example, some issues that would need to be addressed to construct a likelihood for a single pair of daily pocket-meter readings include

- What is the probability distribution of possible pocket-meter readings for a given true dose?
- What is the probability that a single pocket meter or both pocket meters are damaged? (Damaged pocket meters can produce artificially elevated results.)
- What is the probability distribution of possible readings of a damaged pocket meter?
- What rounding and censoring conventions were used?

The computed daily pocket-meter posterior distributions would then be used to compute the cumulative weekly pocket-meter dose distribution, which in turn would be used as the prior $P(x)$ for the weekly film badges. The fact that pocket meters were worn in pairs also provides some interesting possibilities for estimating the above distributions from data rather than just relying on historical information based on pocket-meter measurements.

Another Likelihood Formulation

The likelihood function for $z = 0$, shown in Fig. 6 is based on the censoring convention of recording as zero any reading that would be rounded below 0.3 mSv. When the weekly pocket-meter dose w is zero, it may be reasonable to conclude for $z = 0$ that the expressed dose $\tilde{z}$ was within rounding error of zero and no censoring was necessary. Fig. 10 shows the likelihood function of x, the true dose, when $z = 0$ and no censoring is performed. This is the "rounded" zero likelihood and to the likelihood of Fig. 6 is the "censored" zero likelihood. The posterior distribution of true dose x that results from the "rounded" zero likelihood is compared to the fixed prior in Fig. 11. Comparing Fig. 11 to Fig. 7, shows that the "rounded" zero likelihood puts considerably more posterior mass near zero, effectively excluding dose above 0.15 mSv.

The effect of the "rounded" zero likelihood on $P(x_T|\mathbf{z})$ is demonstrated on persons A and B. The sequence of weekly pocket-meter doses $\mathbf{w}$ for person A

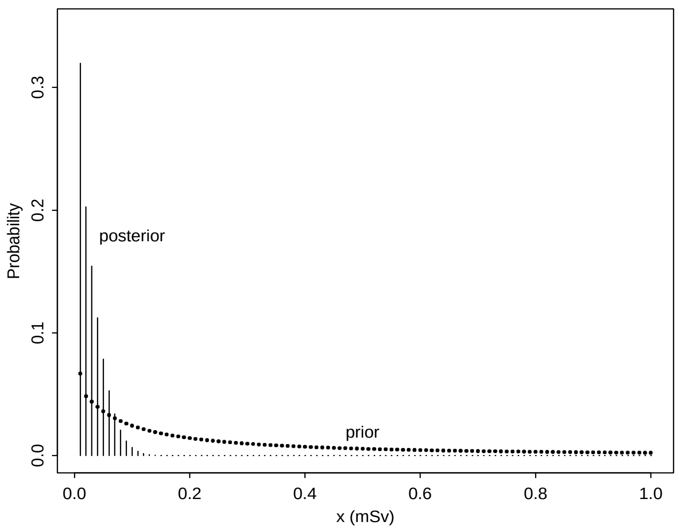

Figure 11: Comparison of prior and posterior distributions of true dose x when recorded dose z is zero without censoring.

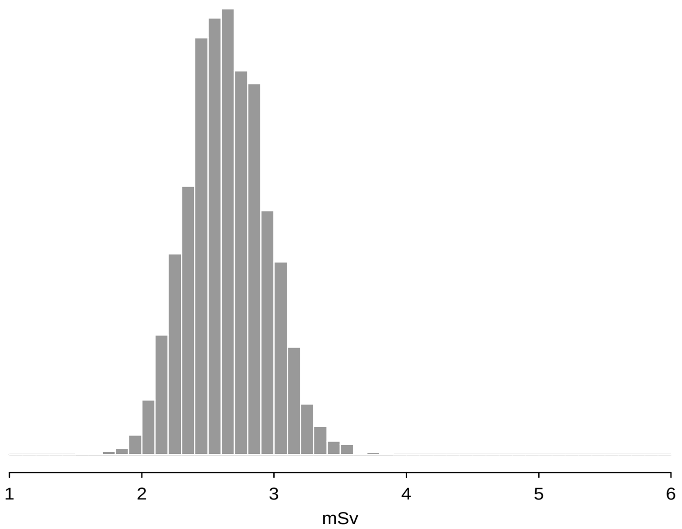

Figure 12: Yearly dose distribution estimate for person A (with recorded dose of 0.65 mSv) allowing "rounded" zeros and with a fixed prior.

is:

$$\{0, 0, 0.05, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.05, 0, 0, 0, 0, 0.05, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0\},$$

and the same sequence for person B is:

$$\{ 0.6, 0.15, 0.85, 0.8, 0.55, 0, 0.5, 0, 0.15, 0.6, 0.55, 4.3, 1.3, 1.7, 1.95, 0.9, 0.45, 1.4, 1.6, 0.4, 0.1, 0.3, 0, 0.05, 0.7, 0, 0.1, 0.1, 0, 0, 0.15, 0, 0, 0.05, 1.7, 0.1, 0, 0, 0, 0.15, 0.35, 0.1, 0.1, 0.7, 0.45, 0.2, 0, 0, 0, 0.35, 0.2, 0.65 \},$$

Their yearly dose distributions are shown in Fig. 12 and Fig. 13. As expected, both distributions have shifted closer to the recorded dose, although both are still centered considerably higher than the recorded dose. The shift toward recorded dose is greater for person A, because of a larger proportion of "rounded" zeros. Rounded zeros also have less uncertainty and this results in a narrower distribution for person A.

A Simple Prior Formulation Based on Pocket-Meter Data

Up to this point, the parameters μ and σ of the prior distribution $P(x)$ have been fixed. To be more objective, the influence of the prior can be reduced by making it more diffuse and some other data can be used in its specification. Pocket-meter data are used to specify μ and σ as follows. The weekly pocket-meter dose, w, is used as the median of $P(x)$ (that is, $\mu = \log(w)$). The parameter σ is set to a value that puts the 95th percentile of $P(x)$ at $w + 6.00$ mSv. This ensures that the prior does not exclude high film-badge readings even when the pocket meter reading is very small or zero. It provides a large amount of prior uncertainty (i.e. the prior is diffuse) and allows the recorded film-badge reading to be the overwhelmingly dominating factor that determines the posterior dose distribution. The posterior distribution is in fact not very sensitive to w due to the large value of σ. More posterior sen-

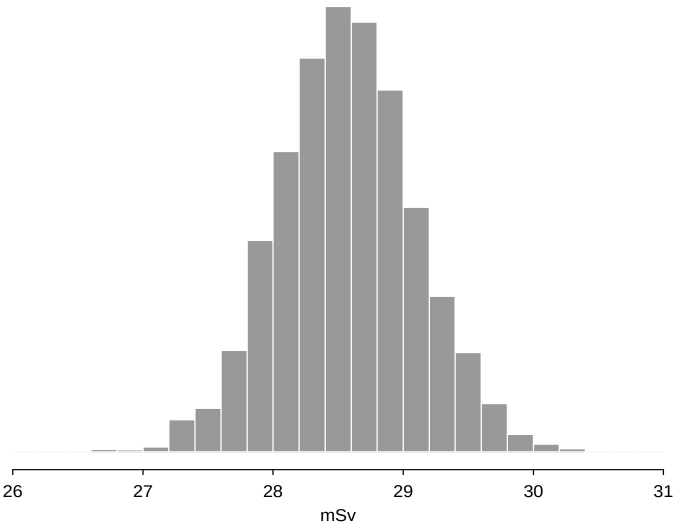

Figure 13: Yearly dose distribution estimate for person B (with recorded dose of 27.2 mSv) allowing "rounded" zeros and with a fixed prior.

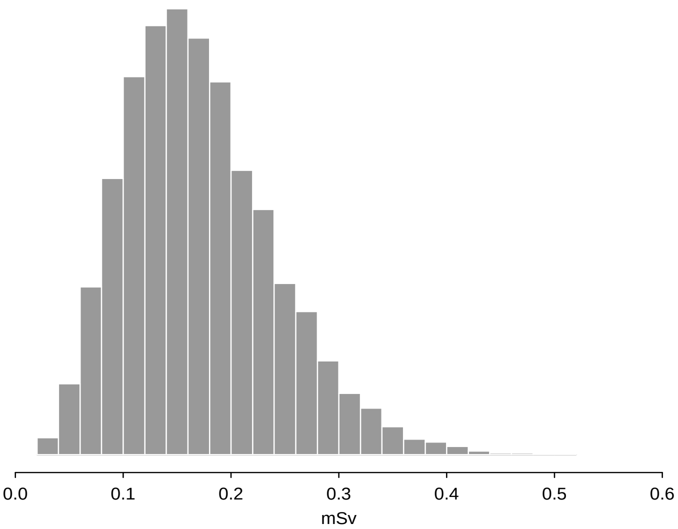

Figure 14: Yearly dose distribution estimate for an individual with zero recorded film-badge reading and all zero weekly pocket-meter doses.

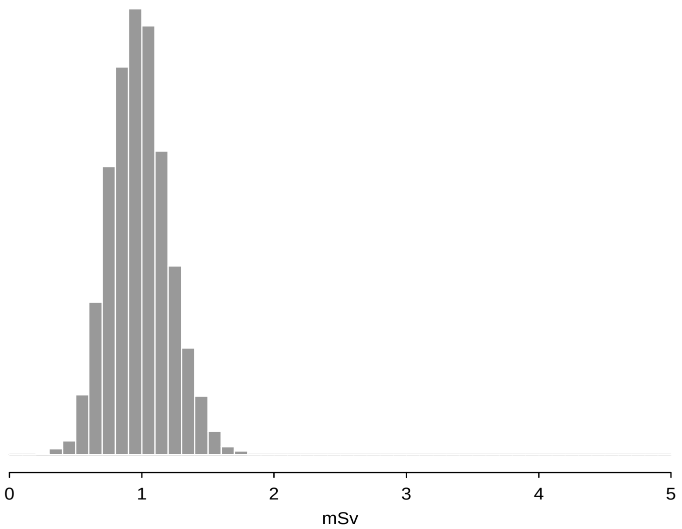

Figure 15: Yearly dose distribution estimate for person A (with recorded dose of 0.65 mSv) allowing "rounded" zeros and with pocket-meter specified priors.

sitivity is exhibited to a constant c_0 discussed later in this section.

Some weekly pocket-meter doses are missing. In such cases that individual's weekly pocket-meter dose average over the weeks that are not missing is used. When all weekly pocket-meter doses are missing, they are treated as zero for the purpose of specifying the prior. (See paragraph below for special treatment of zeros.) This implicitly assumes that an individual with no pocket meter readings is not a radiation worker.

The use of the lognormal distribution as our prior on x, the true dose to the film badge, implicitly assumes that $x > 0$. This is consistent with the belief that the true dose may be negligibly small but can never be zero. Since the pocket-meters have a sensitivity threshold, many weekly pocket-meter doses are recorded as zero. Because the logarithm of zero is undefined, the accepted practice is to choose a small positive value c_0 to replace the zero. The case of $\mathbf{z} = 0$ and $\mathbf{w} = 0$ is used to calibrate the value, since this is where results are most sensitive to it. The value $c_0 = 0.0003$ mSv puts the 95th percentile of $P(x_T|\mathbf{z})$ at 0.30 mSv. That is, c_0 is chosen so that the probability is .95 that the yearly true dose is below 0.30 mSv, when all pocket-meter and film-badge readings are zero. The value 0.30 mSv was chosen, because this is the censoring point for recording zero. This distribution is shown in Fig. 14. As the recorded dose increases, sensitivity of the true dose distribution to the choice of c_0 decreases. For example, the mean of the true dose distribution for an individual with a recorded yearly dose of 16.75 mSv (which consists of about 50% weekly zeros) increases by 0.004 mSv in response to doubling the c_0 value. This is a very small change, particularly considering that c_0 is a weekly quantity and the dose distribution is a yearly quantity.

Fig. 15 and Fig. 16 demonstrate the the effect on $P(x_T|\mathbf{z})$ of specifying the prior parameters μ and σ with pocket-meter data for persons A and B, respectively. Both distribution have again shifted closer to the recorded dose.

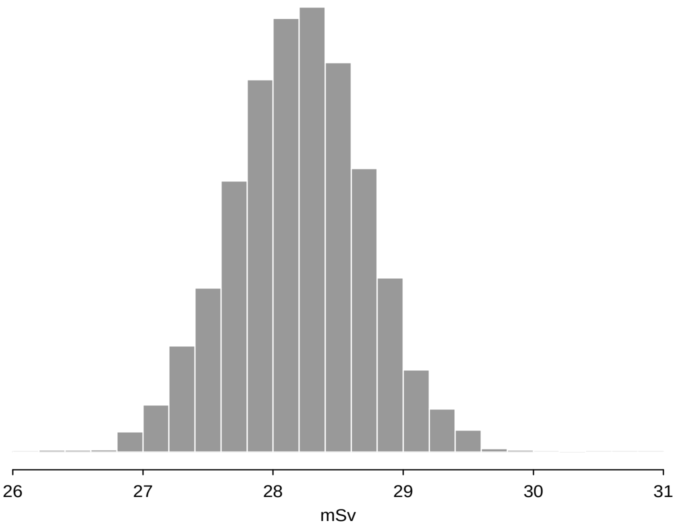

Figure 16: Yearly dose distribution estimate for person B (with recorded dose of 27.2 mSv) allowing "rounded" zeros and with pocket-meter specified priors.

APPLICATION TO *ORNL* COHORT

To illustrate the application of the general dose estimation method and the specific likelihoods and priors developed for the 1945-1955 ORNL cohort, a sample of hard copy records was obtained from the ORNL dosimetry files. The hard copy records contain the detailed daily and weekly monitoring results for each "person-year" selected. The sample was obtained in two stages. It includes a stratified random sample of exposed workers (150 person-years) with yearly film-badge totals that are greater than zero. The remainder of the sample (100 person-years) was obtained by sampling at random from all person-year records over the period from 1945-1955 when film badges were evaluated on a weekly basis. The weekly film-badge reading and the pocket-meter dose were abstracted from the hard copy records. This data has also been used to develop a preliminary dose-adjustment procedure in another study (23). It is important to note that the data currently being used in epidemiologic studies of ORNL workers (9, 24, 3, 8, 25) consist of the yearly total of the weekly film-badge readings for each worker. The detailed weekly records are not available in machine readable form at this time.

Dose Estimates

The dose estimation procedure was applied to the 150 person-years (out of the possible 250 described above) that had at least 30 weekly records. The fact that a person-year does not contain a full 50 weeks (2 weeks are vacation) could be for a number of reasons. We assume that that person worked only the weeks for which there are records. That is, no dose is accumulated for the weeks with no records. Other solutions are possible with additional information. For example, if there is information that a person was assigned to a work area without film-badge monitoring, a dose distribution for this assignment area could be constructed from other data. This distribution would be accumulated for the weeks concerned. The result of this might be a slightly wider yearly dose distribution estimate.

The results are summarized in Figures 17-20, which show boxplots of $P(\mathbf{x}|\mathbf{z})$ for each person-year (labeled with and `id` number and year). The boxplots show the 1, 25, 50, 75, and 99 percentiles of each distribution. In addition, a bold glyph indicates the relative position of the recorded dose for each person-year. The person-

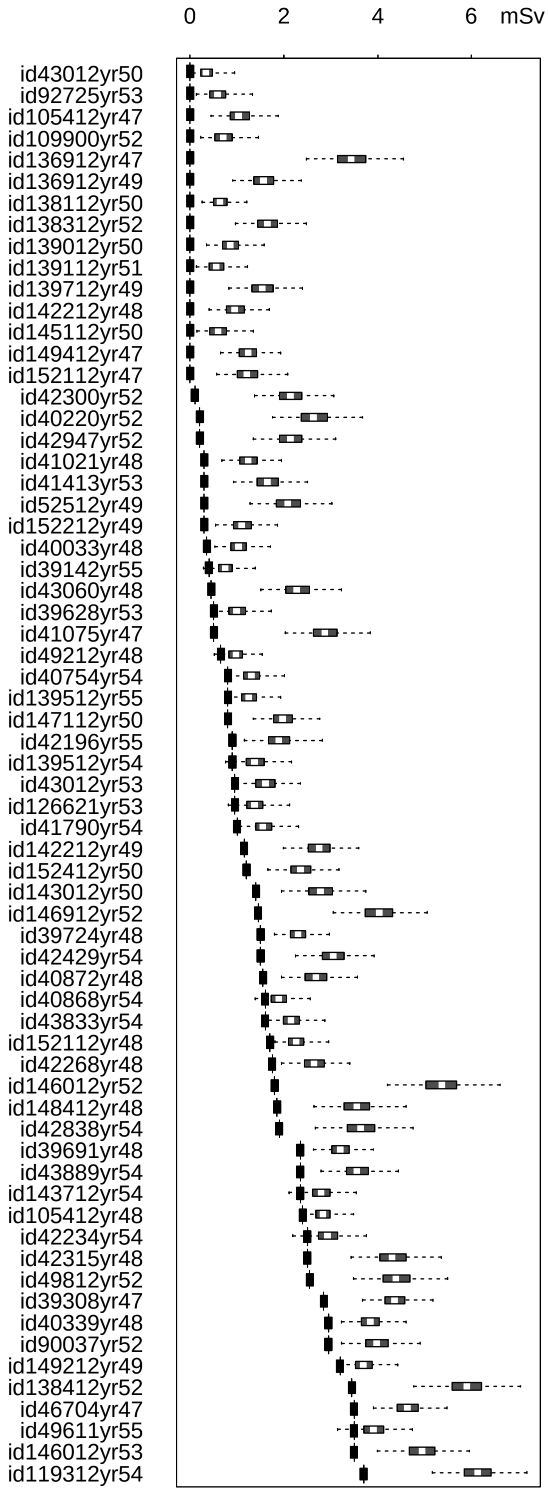

Figure 17: Boxplots of yearly dose distribution estimates and the corresponding recorded dose (mSv) for a sample of the ORNL cohort.

years along the vertical axis are in increasing order of recorded dose. The available pocket-meter data is used to specify priors and modify likelihoods as discussed earlier in this paper. The data contain a few non-

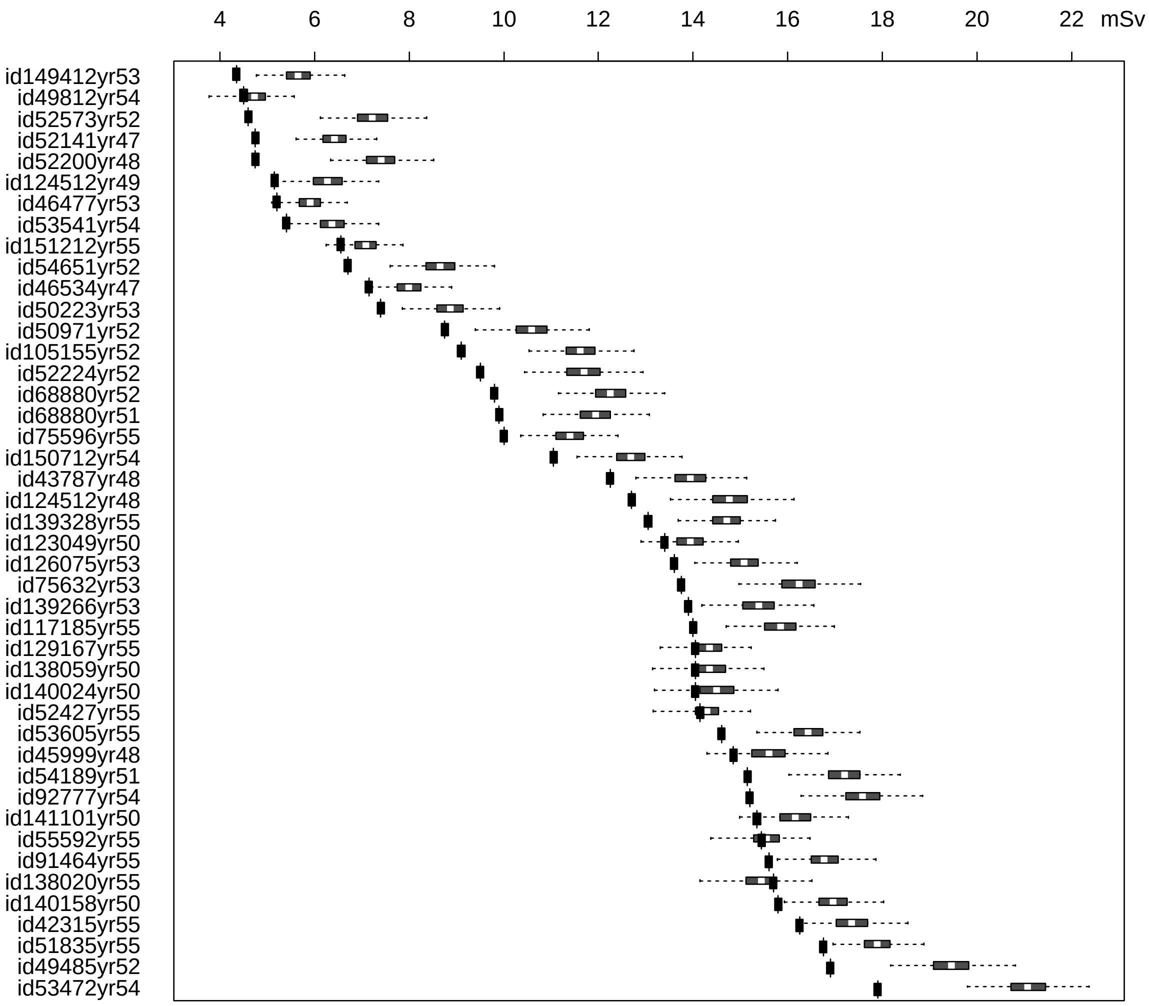

Figure 18: Boxplots of yearly dose distribution estimates and the corresponding recorded dose (mSv) for a sample of the ORNL cohort (continued).

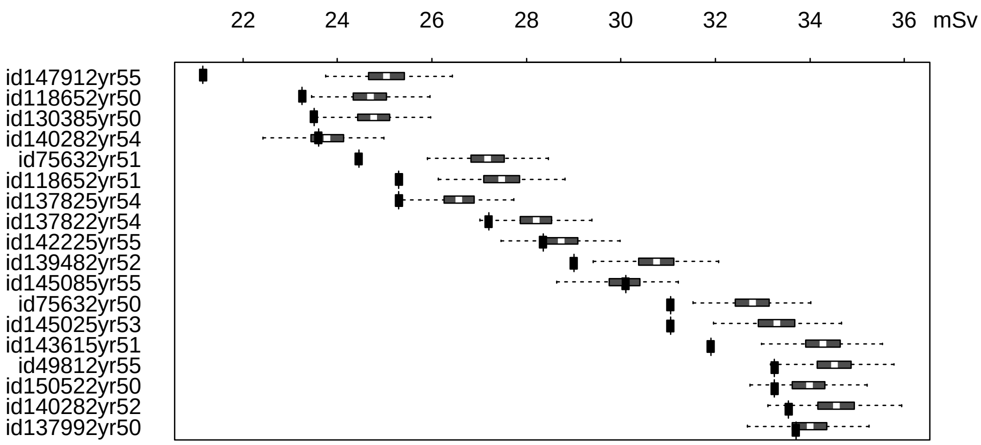

Figure 19: Boxplots of yearly dose distribution estimates and the corresponding recorded dose (mSv) for a sample of the ORNL cohort (continued).

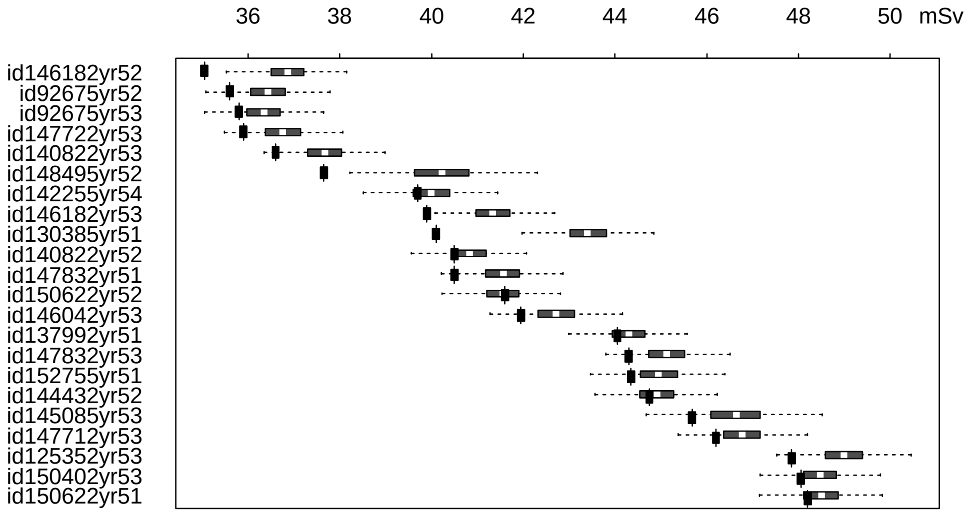

Figure 20: Boxplots of yearly dose distribution estimates and the corresponding recorded dose (mSv) for a sample of the ORNL cohort (continued).

zero recorded film-badge readings below the censoring point of 0.30 mSv. The "rounded" likelihood is used in each of these cases, because clearly no censoring was performed.

The following observations about Figures 17-20 are made with the knowledge of the underlying weekly dose data for each person-year.

- Almost all distributions are centered above the recorded film-badge reading. Thus recorded doses systematically underestimate the true dose.
- The recorded dose is below the one percentile of the distribution in about half the cases. This indicates a severe underestimate of the true dose by the recorded dose.
- The most severe underestimate occurs for person-years that have many zero recorded film-badge readings which correspond to non-zero pocket-meter doses.
- The relative uncertainty is greatest at lower recorded dose levels.
- The large differences in uncertainty at the lower recorded dose levels appear to be mostly due to differences in content of "rounded" and "censored" zeros. Censored zeros ($w > 0$, $z = 0$) introduce more uncertainty than rounded zeros ($w = 0$, $z = 0$).

Dose-Response Analysis

The results indicate that recorded doses for ORNL workers before 1956 that are currently being used in epidemiologic studies contain a large systematic negative bias. Further, there is considerable uncertainty in theses dose estimates that should be taken into account when they are used in dose-response studies of radiation effects. Studies published to date (24, 3, 25, 26) that involve ORNL workers are based on the recorded film-badge readings and have not taken the uncertainties described here into account in the statistical analysis.

A review article (27) discusses several approaches to dose-response analysis mostly from the *structural* (see (21)) view of the measurement error problem. Dose distribution estimates in this report are based on a *functional* view of the measurement error problem. The statistical problem that remains is to develop a specific method for estimating the dose-response parameter given the time to failure and the dose history distributions $P(x(t)|z(t))$ for each individual in the cohort. An obvious, but very computation-intensive solution is simulation.

The practical problem that remains for the ORNL cohort is to obtain the historical daily pocket-meter and weekly film-badge data in electronic form so that valid dose estimates can be obtained for subsequent statistical analysis. Most likely a subset of the data that will support a case control study will be adequate. Thomas, *et al.* (27) report that sample sizes must be several times larger than for the case of no measurement error.

SUMMARY

Methodology is developed to account for uncertainty and bias in measurements of individual occupational radiation dose during 1945-1955 at Oak Ridge National Laboratory (ORNL) using weekly dosimetry data. The product of this methodology is an estimate of the true dose for a person-year in the form of a probability distribution.

The key component of this methodology is a model of the procedures used in a single film-badge measurement. This determines the likelihood function. A realistic model was formulated that includes the most important aspects of film-badge dose measurement in the period 1945 to 1955 at ORNL, however the historical information is often sketchy. The method is very flexible and can easily incorporate much detail. Further, the methodology can be used to asses the sensitivity of dose estimates to various scenarios. The methodology can also be applied to other time periods and other sites by developing new specific models of measurement process.

Another important component is the prior distribution whose importance grows with uncertainty in measurement. Since the largest uncertainty in recorded doses at ORNL results from censoring, dose estimates from data containing many censored zeros are sensitive to prior specification.

It is demonstrated that additional data can be used to specify the prior and modify the likelihood. Pocket-meter data was used to specify the prior location parameter and also to indicate whether a film-badge zero resulted from only rounding or from rounding and censoring. A more rigorous way of including pocket-meter data is by developing a model of the pocket-meter measurement system, applying our methodology to produce a dose estimate from the pocket meters only, and then use it as the prior for the film-badge dose estimation. It would also be possible to use information such as occupation or work location to specify the priors.

The methodology was applied to a sample of dose histories obtained from hard copy dosimetry records at ORNL. The estimated dose distributions show that recorded doses generally have a strong negative bias. The bias is present at all dose levels, but it is most severe at low to medium dose levels, where the recorded dose is usually below the 1 percentile of the true dose estimate. This raises serious questions about the validity of the historical personnel dosimetry data that is used in studies of the effects of low doses in nuclear industry workers. In particular, the results in this report indicate that ORNL workers employed prior to 1957 are likely to have had doses that were higher than those recorded. Consequently, the dose-response coefficients that are based directly on the recorded doses (24, 3, 25, 26) may be biased and their uncertainty is understated.

ACKNOWLEDGMENTS

Research was supported by the Office of Epidemiology and Health Surveillance, Environment, Safety and Health, U. S. Department of Energy under contract DE-AC05-96OR22464. We thank the Office of Radiation Protection at ORNL for providing the data used in this study, and the Center for Epidemiologic Research at Oak Ridge Institute for Science and Education for help with data entry and verification. We also thank Asim YarKhan of the University of Tennessee for his help in implementing some of the software. Finally, we thank two anonymous referees for many useful suggestions that helped us prepare this revision.

References

1. J. C. Hart. A progress report dealing with the derivation of dose data from ORNL personnel exposure records applicable to the Mancuso study. Technical Report ORNL/M-2614 (formerly, ORNL CFM 66-1-84), Oak Ridge National Laboratory, Oak Ridge, TN, 37831, January 14 1966.

2. R. H. Wilson, W. V. Baumgartner, J. J. Fix, and L. L. Nichols. Description and evaluation of the hanford personnel dosimeter program from 1944 to 1989. Technical Report PNL-7447, Pacific Northwest Laboratory, Richland, WA, 1990.

3. S. Wing, C. M. Shy, J. L. Wood, S. Wolf, D. L. Cagle, and E. L. Frome. Mortality among workers at Oak Ridge National Laboratory: Evidence of radiation effects in follow-up through 1984. *JAMA*, 265:1397–1402, 1992.

4. George D. Kerr. Missing dose from mortality studies of radiation effects among workers at Oak Ridge National Laboratory. *Health Phys.*, 66:206–208, 1994.

5. G. W. Kneale, T. Sorahan, and A. M. Stewart. Evidence of biased recording of radiation doses of Hanford workers. *Am. J. Ind. Med.*, 20:799–803, 1991.

6. Fred C. Maienschein and R. W. Peelle. Radiation dosage estimation and health risk. *JAMA*, 267:929, 1992.

7. S. Wing, C. M. West, J. L. Wood, and W. Tankersley. Recording of external radiation exposures at Oak Ridge National Laboratory: Implications for epidemiologic studies. *J. Exposure Analysis and Environmental Epidemiology*, 4(1):83–93, 1994.

8. S. Wing, C. M. Shy, J. L. Wood, S. Wolf, and D. L. Cragle. Reply to comments by Maienschein and Peelle. *JAMA*, 267:929–930, 1992.

9. E. Cardis and J. Estéve. Uncertainties in recorded doses in the nuclear industry: identification, quantification and implications for epidemiological studies. *Radiat. Prot. Dosim.*, 36(2/4):315–319, 1991.

10. E. S. Gilbert. Accounting for bias and uncertainty resulting from dose measurement errors and other factors. Technical Report PNL-SA-18907, Pacific Northwest Laboratory, Richland, WA 99352, February 1990.

11. Daniel J. Strom. Estimating individual and collective doses to groups with 'less than detectable' doses: A method for use in epidemiologic studies. *Health Phys.*, 51(4):437–445, 1986.

12. N. A. Taylor. Estimation of dose received when dosemeter results are recorded below a threshold level. *J. Radiol. Prot.*, 11(3):191–198, 1991.

13. K. Z. Morgan. Dosimetry requirements for protection from ionizing radiation. In *Selected Topics in Radiation Dosimetry*, pages 3–23. Vienna: International Atomic Energy Agency, IAEA Proceedings Series, 1961.

14. K. Z. Morgan. Techniques of personnel monitoring and radiation surveying. In A. H. Snell, editor, *Nuclear Instruments and Their Uses*, volume 1, pages 391–469. John Wiley & Sons, 1962.

15. E.S. Gilbert, J.J. Fix, and W.V. Baumgartner. An approach to evaluating bias and uncertainty in estimates of external dose obtained from personal dosimeters. *Health Phys.*, 70(3):336–345, 1996.

16. Toby J. Mitchell, George Ostrouchov, Edward L. Frome, and George D. Kerr. A method for estimating occupational radiation dose to individuals, using weekly dosimetry data. Technical Report ORNL-6778, Oak Ridge National Laboratory, Oak Ridge, TN 37831, December 1993.

17. George Ostrouchov, Asim YarKhan, and Edward L. Frome. Computing occupational radiation dose distributions for individuals. Technical Report ORNL-6792, Oak Ridge National Laboratory, Oak Ridge, TN 37831, 1997.

18. James S. Press. *Bayesian Statistics: Principles, Models, and Applications.* John Wiley & Sons, New York, 1989.

19. G. Miller, W. C. Inkret, and H. F. Martz. Bayesian detection analysis for radiation exposure. *Radiation Protection Dosimetry*, 48(3):251–256, 1993.

20. G. Miller, W. C. Inkret, and H. F. Martz. Bayesian detection analysis for radiation exposure, II. *Radiation Protection Dosimetry*, 58(2):115–125, 1995.

21. Wayne A. Fuller. *Measurement error models.* John Wiley & Sons, Inc., New York, 1987.

22. William J. Kennedy, Jr. and James E. Gentle. *Statistical Computing.* Marcel Dekker, New York, 1980.

23. E. L. Frome, D. L. Cragle, J. Watkins, S. Wing, C. Shy, W. Tankersley, and C. M. West. A mortality study of employees of the nuclear industry in Oak Ridge, Tennessee. *Radiat. Res.*, in press. (also ORNL-6785 Technical Report).

24. E. S. Gilbert, S. A. Fry, L. D. Wiggs, G. L. Volez, D. L. Cragle, and G. R. Petersen. Analyses of combined mortality data on workers at the Hanford Site, Oak Ridge National Laboratory, and Rocky Flats Nuclear Weapons Plant. *Radiat. Res.*, 120:19–35, 1989.

25. Steve Wing, Carl M. Shy, Joy L. Wood, Susanne Wolf, Donna L. Cragle, William Tankersley, and E. L. Frome. Job factors, radiation and cancer mortality at Oak Ridge National Laboratory: Follow-up through 1984. *Am. J. Ind. Med.*, 23:265–279, 1993.

26. E. Cardis, E.S. Gilbert, L. Carpenter, G. Howe, I. Kato, B.K. Armstrong, V. Beral, G. Cowper, A. Douglas, J. Fix, S.A. Fry, J. Kaldor, C. Lave, L. Salmon, P.G. Smith, G.L. Voelz, and L.D. Wiggs. Effects of low doses and low dose rates of external ionizing radiation: cancer mortality among nuclear industry workers in three countries. *Radiat. Res.*, 142:117–132, 1995.

27. Duncan Thomas, Daniel Stram, and James Dwyer. EXPOSURE MEASUREMENT ERROR: influence on exposure-disease relationships and methods of correction. *Annu. Rev. Publ. Health*, 14:69–93, 1993.