RADIATION RESEARCH **170**, 118–126 (2008)
0033-7587/08 $15.00

Allowance for Random Dose Estimation Errors in Atomic Bomb Survivor Studies: A Revision

Donald A. Pierce,[a,1] Michael Væth[b] and John B. Cologne[c]

[a] *Division of Biostatistics, Department of Public Health, Oregon Health Sciences University, Portland, Oregon;* [b] *Department of Biostatistics, Aarhus University, Aarhus, Denmark; and* [c] *Department of Statistics, Radiation Effects Research Foundation, Hiroshima, Japan*

Pierce, D. A., Væth, M. and Cologne, J. B. Allowance for Random Dose Estimation Errors in Atomic Bomb Survivor Studies: A Revision. *Radiat. Res*. 170, 118–126 (2008).

Allowing for imprecision of radiation dose estimates for A-bomb survivors followed up by the Radiation Effects Research Foundation can be improved through recent statistical methodology. Since the entire RERF dosimetry system has recently been revised, it is timely to reconsider this. We have found that the dosimetry revision itself does not warrant changes in these methods but that the new methodology does. In addition to assumptions regarding the form and magnitude of dose estimation errors, previous and current methods involve the apparent distribution of true doses in the cohort. New formulas give results conveniently and explicitly in terms of these inputs. Further, it is now possible to use assumptions about two components of the dose errors, referred to in the statistical literature as "classical" and "Berkson-type". There are indirect statistical indications, involving non-cancer biological effects, that errors may be somewhat larger than assumed before, in line with recommendations made here. Inevitably, methods must rely on uncertain assumptions about the magnitude of dose errors, and it is comforting to find that, within the range of plausibility, eventual cancer risk estimates are not very sensitive to these.

INTRODUCTION

For analyses of the A-bomb survivor data, the Radiation Effects Research Foundation (RERF) has for more than 15 years made adjustments to reduce the effects of imprecision in radiation dose estimates in the manner developed in refs. (*1, 2*). More recently, general technical progress has been made (*3*) for carrying out such adjustments, and the aim here is to apply this to the RERF setting. A thoroughly revised dosimetry system DS02 (*4*) has recently been implemented to replace the previous DS86, making it timely to reconsider these matters. All results in this paper are in terms of DS02. We have found that the dosimetry revision itself does not warrant significant changes in the adjustment methods, but we recommend two changes for other reasons. The first is to use the general formulas developed in ref. (*3,*) as explained here. The second is to modify the assumptions regarding the magnitude of dose estimation errors in view of matters more well understood now than in refs. (*1, 2*). These will not result in appreciable changes in overall risk estimation, although it is possible that there could be greater effects on more specialized investigations. At any rate, it is advisable to have the ingredients for the adjustments, considered here, as accurate as possible.

It is well known that dose estimation errors that are not systematic but vary independently among survivors, although largely averaging out in fitting dose–response models, cause some systematic downward bias in risk estimation; see e.g. refs. (*5–7*). The size of this bias depends in complicated ways on the particular setting, but for the RERF data we can deduce that individual dose estimation errors, when assumed to be typically in the range 35–50%, result in approximately a 10–15% downward bias in radiation risk estimates. This level of bias would not be very important in relation to other uncertainties, but it is important to confirm that it is not substantially larger than this and to continue taking measures to reduce it. The approach discussed below is in principle a standard one (*5*), but as indicated there are some special features involved for the A-bomb survivor cohort.

There are also systematic errors affecting large groups of survivors that are not considered in this paper but are worthy of further attention. These include the yields of the two bombs and technical approaches used in the dosimetry system to account for individual shielding. Some analysis of these is made in the final chapter of ref. (*4*), to which we refer again later in this paper.

First we review the basic issues involved for random dosimetry errors. Simplifying some in regard to aspects of dose, we refer to unknown true doses by x, and estimates from the dosimetry system by z. The basic aim of methods developed in refs. (*1, 2*) is to replace z in dose–response analyses by *adjusted dose estimates* $E(x|z)$, the expected value of true dose given the estimated dose. It will surprise many that the aim of a dosimetry system is not that $E(x|z)$

[1] Address for correspondence: Division of Biostatistics, Department of Public Health, 3181 SW Sam Jackson Park Road, Mail code CR145, Portland, OR 97239-3098; e-mail: pierce.don.a@gmail.com.

$= z$, so discussion of that is warranted. The aim of a dosimetry system is essentially that $E(z|x) = x$, although this may be better considered as holding on a log scale. Even when the dosimetry system estimates are unbiased in this sense, this does not mean that survivors in the study cohort will have $E(x|z) = z$, and indeed in the higher dose range it is inevitable that $E(x|z) < z$; see also ref. (*1*) for elaboration on that.

This will fairly generally be the case whenever the cohort distribution of true doses is skewed with a long tail to the right, but there are useful considerations specific to the A-bomb survivors. The dosimetry system uses estimates of survivor location and shielding to arrive at dose estimates. It does not aim to use the information provided by the fact of a person's *survival*, which certainly carries some information about true dose, and we can say that the aim is largely to incorporate this further information. That is, the dosimetry system could in principle be applied not just to survivors but to all exposed persons. The condition $E(z|x) = x$ does not depend on whether one is considering all exposed persons or only survivors. However, the value of $E(x|z)$ will differ between those frames of reference, and we are interested in the value of this among survivors. Among survivors with a given dose estimate z, there is a range of true doses with considerably more survivors in the lower part of that range than in the upper part, because survival was strongly related to true dose. The result is that except for quite low dose estimates where survival was not much of an issue, it must be true that among survivors $E(x|z) < z$. The primary oversimplification in this argument is that the frame of reference to be considered is not actually survivors, but those comprising the study cohort; this would be the appropriate frame of reference for defining $E(x|z)$ for any study.

Ideally speaking, fitting models in terms of $E(x|z)$ rather than z will eliminate most of the bias due to dose estimation errors when the dose response is linear in x. [For models involving x^2 the same is achieved by substituting for this $E(x^2|z)$ rather than z^2]. However, the estimation of $E(x|z)$ for the survivor frame of reference is technically quite difficult. Even though the principles are reasonably clear, the implementation of them in ref. (*1*) was quite cumbersome and approximate, as explained in ref. (*3*). What has emerged in our recent research (*3*), and is the main point here, is a much more convenient and accurate way of carrying this out.

Moreover, the recent results in ref. (*3*) allow for dealing with an issue that was not addressed before. There are two types of random errors inherent in the dosimetry system. The type dealt with before, which we will call *measurement* errors, result basically from random errors in estimating the individual survivor location and shielding situation. The other type that can now be dealt with we will refer to as *averaging* errors. These arise at least in part from explicit grouping, averaging and use of smoothing formulae in the DS02 treatment of location and shielding. These errors might remain even if location and shielding were known exactly. The two types of errors have very different effects on risk estimation, and with the new methods developed in ref. (*3*) and reviewed here we can allow for both to be present.

RECOMMENDED NEW METHOD

We allow for a combination of what we call measurement and averaging errors, whose distinction is indicated above. In the literature on covariate errors these are usually called, respectively, ''classical'' and ''Berkson'' errors. As indicated in ref. (*5*), there is much literature on this, and we point especially to the clear development in ref. (*6*) of basic methods and to ref. (*7*) discussing the matter specifically for radiation studies.

Technically, measurement errors are those on a suitable scale uncorrelated with the true values, whereas averaging errors are uncorrelated with the estimated values, as would be a consequence of regression smoothing. As in refs. (*1–3*), we assume that the measurement errors are on a log scale normally distributed with standard deviation σ_M, and that $E(\log z|x) = \log x$. That is, on the dose scale σ_M is approximately the coefficient of variation resulting from this aspect of dose errors. We further assume that the averaging errors contribute on the log scale normally distributed error with standard deviation σ_A, with similar coefficient-of-variation interpretation. Basically, the model employed in ref. (*3*) is that $\log(z) = \log(x) + e_M + e_A$, with e_M uncorrelated with x, e_A uncorrelated with z, e_M uncorrelated with e_A , and with (σ_M, σ_A), respectively, the standard deviations of (e_M, e_A). These relations were developed in ref. (*3*), are equivalent to the formulation used in ref. (*6*) provided that the distributions of (e_M, e_A) are symmetric, and through the correlation structure employed may show more clearly the fundamental distinction between the two types of error. The desired adjustments depend on the two (log scale) variances and also on further aspects of the distribution of true doses in the study cohort.

It is convenient, as done previously, to express the desired estimate of $E(x|z)$ in terms of an adjustment factor applied to the estimated doses, so that $E(x|z) = C(z)z$. For reasons that will become apparent, this adjustment factor differs between Hiroshima and Nagasaki. Some readers may want to skip the technicalities of the next few paragraphs and proceed directly to consideration of Fig. 1. The basis of the new method developed in ref. (*3*) is an excellent approximation of form

$$\log[C(z)] = \frac{1 + 2d_1(z)}{1 - \sigma_M^2 d_2(z)} \frac{\sigma_M^2}{2}, \qquad (1)$$

where $d_j(z)$, $j = 1, 2$ are the first two derivatives on a log-log scale of the density function of true doses, evaluated at $x = z$. The value of σ_A does not enter into this approximation, although it must be used as indicated below. The point of Eq. (1) is to allow for smooth departures of the

density of $\log(x)$ from the normal case, representing this departure in terms of a smoothly varying $d_2(z)$. Related methods for approximating $E(x^2|z)$, important for several reasons, are discussed later.

When $\log(x)$, as well as the errors, is normally distributed, then $d_2(z)$ is constant and the value of $C(z)$ resulting from Eq. (1) becomes the well-known result for that lognormal setting. A version of this result found in Eq. (12) of ref. (*7*) is that $\log E(x|z) = (1 - R^2) \log E(x) + R^2 \log z$, where $R^2 = \sigma^2/(\sigma^2 + \sigma_M^2)$ with $\sigma^2 = \mathrm{var}(x)$. It is easily calculated from both this result, and Eq. (1) for the lognormal case, that $\log E(x|z) = (\sigma_M^2\sigma^2/2)/(\sigma^2 + \sigma_M^2) + (1 - R^2)E(\log x) + R^2 \log z$. This is a weighted average of $E(\log x)$ and $\log z$, plus a constant term that adjusts for the relation $\log E(x)/E(\log x) = \sigma^2/2$. The form of this constant term arises from our assumption that $E(\log z|x) = \log x$ rather than $E(z|x) = x$. Use of the latter assumption is also considered in ref. (*3*), where it is shown that in that case Eq. (1) is modified by adding $\sigma_M^2/2$ to $\log\{C(z)\}$. This change is numerically substantial, pointing to the importance of which assumption is made, which is emphasized as well in ref. (*7*).

Since the distribution of true doses is not observed, we must first approximate the required derivatives in terms of corresponding derivatives of the distribution of observed doses. Write $\hat{d}_1(z)$ and $\hat{d}_2(z)$ for the first two log-log derivatives of the distribution of estimated doses. These derivatives can be estimated very well for the RERF cohort by approximating the log-log density with second- and third-degree polynomials. Specifics of this analysis, and resulting formulas for each of Hiroshima and Nagasaki, are given in the Appendix. As was shown in ref. (*3*), under the idealized modeling indicated in the second paragraph of this section, the corresponding derivatives of the distribution of true doses, evaluated at $x = z$, required for implementing Eq. (1), are given approximately by

$$d_1(z) = \hat{d}_1(z)/[1 + (\sigma_M^2 - \sigma_A^2)\hat{d}_2(z)]$$
$$d_2(z) = \hat{d}_2(z)/[1 + (\sigma_M^2 - \sigma_A^2)\hat{d}_2(z)]. \qquad (2)$$

Use of these relations circumvents what was the most clumsy part of the approach in refs. (*1, 2*), sometimes referred to as ''deconvolution''. Surprisingly, the dominant effect of increasing σ_A is to decrease $d_1(z)$ as given by Eq. (2), and hence to result in less adjustment at high doses.

For these considerations one should set aside the zero-dose (i.e. $z < 0.005$ Gy) portion of the cohort, assuming that if $z = 0$ then $x = 0$ as well. (Most of those with $z = 0$ were too far from the bombs to have $x > 0$). In fact, the rationale for the approximations breaks down for sufficiently small dose estimates, and for $z < 0.2$ Sv one should take $C(z) = C(0.2)$. Otherwise, the approximations (1) and (2) are excellent provided that $d_2(z)$ does not vary too rapidly, which it does not for the RERF cohort. The main feature of the adjustments is that $C(z) < 1$ for large doses, but as seen here, the values of $C(z)$ are greater than unity for parts of the low-dose range that differ between Hiroshima and Nagasaki. Such adjustments were not used in the past but should be now with the more accurate methods.

The methods used before arrived at the adjustment factors $C(z)$ by much more cumbersome and inaccurate methods described in ref. (*3*). The assumptions made previously were as above, with $\sigma_M = 0.35$, $\sigma_A = 0$. Although all calculations in this paper use DS02, we note that the functions $\hat{d}_j(z)$, $j = 1, 2$ (not used in previous implementation) are essentially the same for the DS86 and DS02 dose estimates. This is because those functions are unchanged by an overall rescaling of dose estimates and, in terms of the effect on the frequency distribution of dose estimates, the dosimetry revision is very close to that rescaling. Thus differences between adjustment factors $C(z)$ computed before and those computed by present methods taking $\sigma_M = 0.35$, $\sigma_A = 0$ are due to inadequacies in the previous calculations, which is not surprising in view of the methods used before.

Figure 1 gives some examples of the functions $C(z)$ for each city, for selected values of the parameters (σ_M, σ_A), with comparison to the adjustment factors that have been used until now. The rationale for the choices of (σ_M, σ_A) pertains to currently made adjustments and those proposed in this paper; effects of these on risk estimation will be demonstrated later. The values of percentage error, i.e. the approximate coefficient of variation of dose estimates, for that figure refer directly to values of (σ_M, σ_A). We see that the inadequacies of previous calculations were particularly notable for Nagasaki. This is because the ''shape'' of the distribution of x is somewhat different for Nagasaki than for Hiroshima, whereas in the previous calculations these were taken as the same. In particular, the distribution of $\log(x)$ is considerably closer to a normal distribution for Hiroshima than for Nagasaki. City differences in this respect are mainly due to two factors: the Nagasaki bomb being much more powerful than the Hiroshima one and the geographic distribution of survivors being much more circular for Hiroshima.

SOME DETAILS OF RECOMMENDED IMPLEMENTATION

We now present some aspects of recommended implementation that are independent of the values of (σ_M, σ_A), to be used. These recommendations correspond essentially to the previous RERF implementation, although not all are discussed in refs. (*1, 2*).

First, adjustments are required for DS02 doses that are provided for major organs and for both the γ-ray and neutron components of these. We recommend that correction factors $C(z)$ be computed in terms of total shielded kerma (not organ-specific and used as γ-ray plus neutron) and that these factors be applied to each organ dose. They may also be applied to each of the γ-ray and neutron components when these are required. Alternatives to this would be hopelessly complicated to implement and would go far beyond the aims of the adjustments.

Second, since the dose response for most effects has

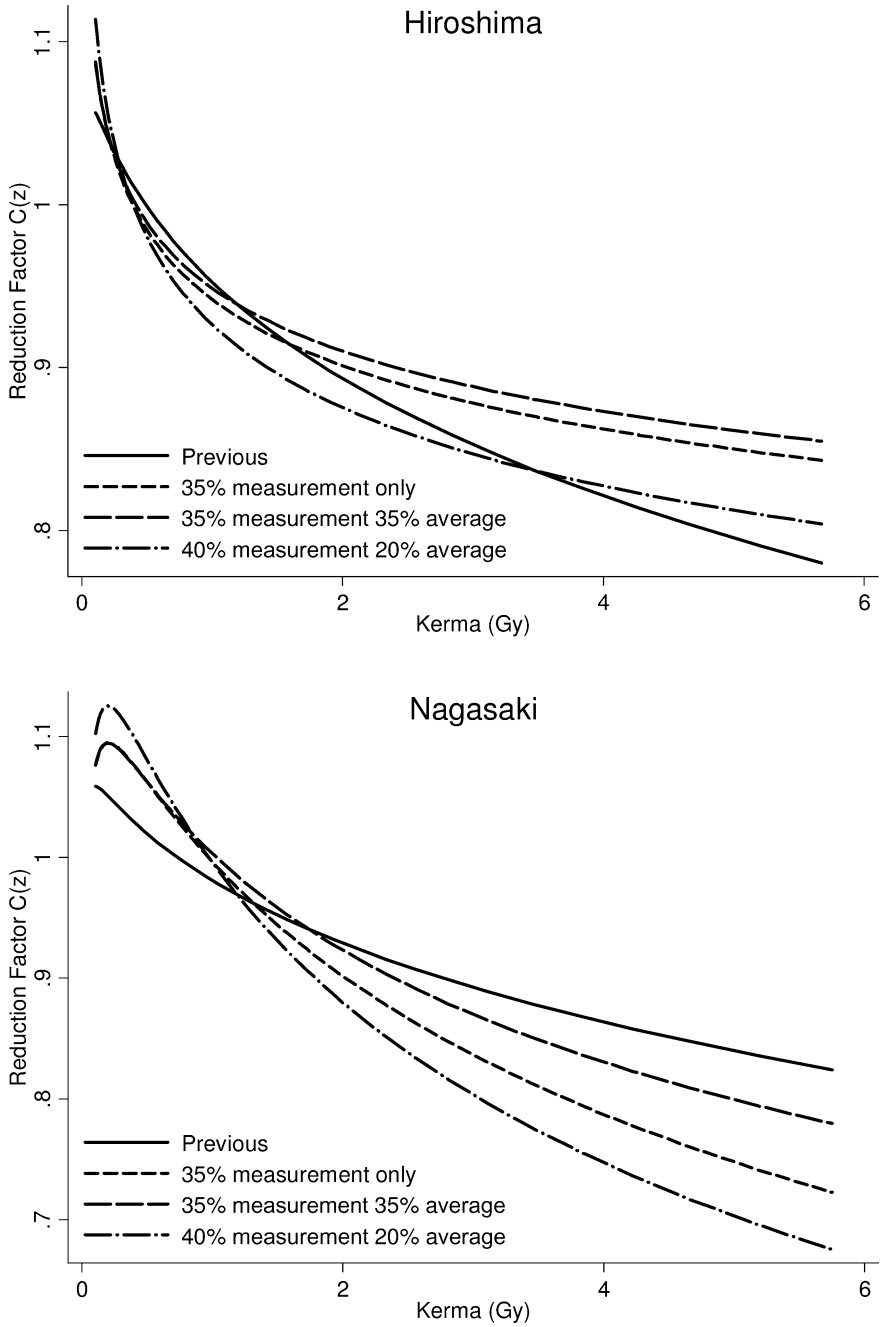

FIG. 1. Examples of adjustment factors $C(z)$ for various values of σ_M and σ_A written here as the percentage errors, along with those for the previous (current) method. The differences between ''previous'' and ''35% measurement only'' are due to inadequacies in previous calculations. Note that incorporating averaging-type errors results in less adjustment. Risk estimation is mainly affected by the adjustments at moderately high doses, e.g. >2 Gy kerma.

some leveling off at high doses, it is usually appropriate for routine analyses of dose response effectively to omit some small fraction of the highest-dose survivors. Some examination of this, and justification of what we recommend, is given below. However, for analyses of sex-age-time patterns of risk it is important to use the highest-dose survivors. Truncating the highest doses rather than omitting those with such doses can be used to avoid using different dose ranges for these two purposes. This also serves to avoid overdependence on the dose error model at the highest doses. Truncation should be done in terms of shielded kerma, as follows, and results can be used for both purposes considered above. For the moment we assume this truncation is done at 4 Gy kerma and note that, of the 86,661 survivors with known doses used for cancer mortality analyses, there are only 323 with kerma level over 4 Gy.

The recommended procedure for each city, for any chosen values of the parameters (σ_M, σ_A), is as follows. Letting z denote shielded kerma in Gy, write the truncated version as $z_t = \min(z, 4)$. Compute the adjustment factor discussed

above as $C(z_t)$, so that its value for kerma values greater than 4 is not used. Then, for any organ dose d_o, not truncated, and for each of the γ-ray and neutron components, compute the adjusted organ dose as

$$d_{oa} = \frac{z_t}{z} C(z_t) d_o. \quad (3)$$

We will illustrate the effect of this 4 Gy truncation by comparison to truncation in the same manner at 6 Gy (some truncation is necessary since the highest kerma estimates are over 50 Gy). Adjusted doses for this purpose are taken for the case $\sigma_M = 0.40$, $\sigma_A = 0.20$, which is recommended in the final section. The top panel in Fig. 2 illustrates the estimated dose response for solid cancer mortality, in terms of (unadjusted) kerma categories, with the highest category being >6 Gy. Results for *adjusted* doses with 6 Gy truncation are shown in the middle panel, where the same dose category-specific ERRs as in the top panel are plotted as a function of the adjusted truncated dose. The dose adjustment progressively reduces the values of dose class-marks, but not enough to alter the basic shape of the dose response. The bottom panel gives results for adjusted doses with 4 Gy truncation in the manner described in Eq. (3). Again the ERRs for the kerma categories are plotted as a function of adjusted dose, but now all the kerma categories over 4 Gy have about the same adjusted truncated dose, so those categories are combined. Clearly, if adjusted doses with 6 Gy truncation were used, some special allowance would be needed for the higher dose categories. This is not necessary with 4 Gy truncation, allowing for unified analysis of both dose response and age-time-sex effects, and this is employed in most RERF reports.

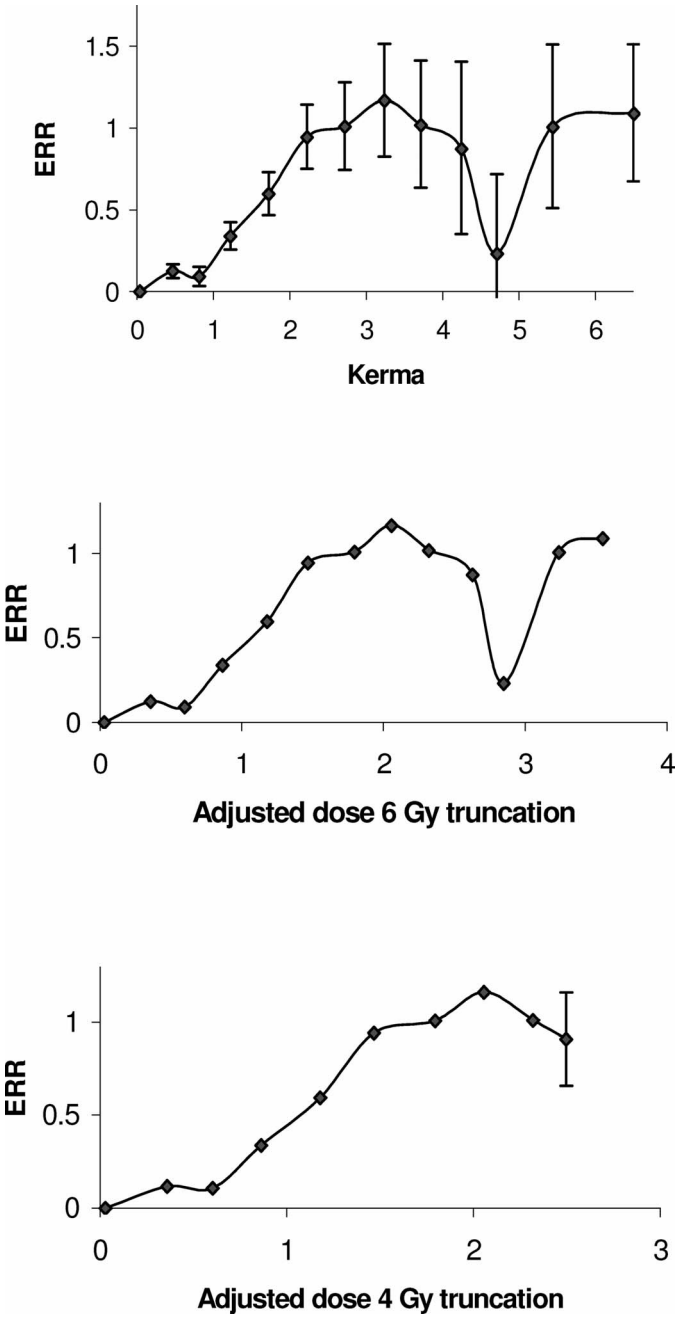

FIG. 2. Solid cancer mortality ERR as a function of kerma and two versions of adjusted dose when taking $\sigma_M = 0.4$, $\sigma_A = 0.2$. In the lower two panels, dose is that to the colon, roughly 60% of kerma. Error bars represent one standard error, and except for the one shown in the bottom panel they are the same for all panels. ERR is presented for exposure age 30, averaged over sex.

SENSITIVITY OF RISK ESTIMATES TO ASSUMED ERROR LEVELS

For the choices of adjustment shown in Fig. 1, we illustrate in Table 1 the effect of the adjustments for a primary setting of risk estimation, namely for all solid cancers together based on mortality data with follow-up through 2000. These data are publicly available at http://rerf.or.jp (DS02 Risk Estimation, Solid Cancer and Leukemia Mortality Data), documented further in ref. (*8*). For simplicity we use here the age-constant ERR model of ref. (*9*), presenting the ERR/Sv for exposure age 30, averaged over sex. Aside from the primary risk estimate, some attention is given to inferences about city differences and downward curvature of the dose response when restricting the dose range to avoid the leveling off of the dose response. The model used for this analysis is essentially ERR = β dose $\exp\{\gamma_1 \text{city} + \gamma_2 \text{sex} + \gamma_3 agex\}$, where variables are coded so that β is standardized as indicated above (*agex* denotes exposure age).

Although it is clear from Fig. 2 that linear risk estimation on the full dose range should not be used without the truncation corresponding to 4 Gy kerma, we also give results for truncation to 6 Gy kerma. The reason for this is only to indicate how much of the change in estimates is due to the reduction factors shown in Fig. 1 and how much is due to the truncation.

The ERRs for exposure ages other than 30 years are computed by adding to the logarithm of the ERR shown the product of (*agex* − 30) in decades and the parameter estimates shown under *exp age*. The *city* effect shown is the ratio of estimated ERRs for Hiroshima and Nagasaki, with the *P* value for testing no city difference given in

TABLE 1
Results of Various Adjustments as Shown in Fig. 1 on Risk Estimation for Solid Cancer Mortality

Method	ERR (30)	Increase	Exposure age	City (*P*)	Curvature (*P*)
Truncation at 4 Gy kerma					
Unadjusted	0.42		−0.38	1.14 (0.14)	0.56 (0.02)
Previous adjustment	0.47	1.12	−0.38	1.18 (0.09)	1.10 (0.007)
35% measurement (new)	0.47	1.12	−0.38	1.16 (0.12)	1.11 (0.007)
35% measurement 35% average	0.46	1.10	−0.38	1.17 (0.11)	1.04 (0.008)
40% measurement 20% average	0.48	1.14	−0.38	1.17 (0.11)	1.28 (0.005)
Truncation at 6 Gy kerma					
Unadjusted	0.39		−0.40	1.15 (0.12)	0.56 (0.02)
Previous adjustment	0.44	1.13	−0.38	1.17 (0.09)	1.10 (0.007)
35% measurement (new)	0.44	1.13	−0.39	1.14 (0.15)	1.11 (0.007)
35% measurement 35% average	0.43	1.10	−0.39	1.15 (0.12)	1.04 (0.008)
40% measurement 20% average	0.45	1.15	−0.38	1.15 (0.13)	1.28 (0.005)

Notes. ERR (30) is the solid cancer excess relative for exposure age 30, averaged over sex, and ''increase'' is the ERR modification factor resulting from dose adjustments. Remaining columns pertain to exposure age, city effect and curvature estimates. *P* values are for testing null hypotheses of no city effect and no curvature.

parentheses. The parameter estimate under *Curvature* is the ratio of the quadratic to linear coefficients in what is called a linear-quadratic (LQ) model, with the *P* value for testing that this is zero given in parentheses.

An LQ model takes the ERR as $1 + \beta x + \gamma x^2$, with the curvature defined as γ/β, and fitting such models requires an approximation to $E(x^2|z)$. Approximations for this are also developed in ref. (*3*) using results similar to that in Eq. (1). For our purposes it suffices to note that a very good approximation, $E(x^2|z) = 1.15\{E(x|z)\}^2$ for the RERF data, in both cities and for all values of σ_M, σ_A that might reasonably be used.

In terms of the ERR estimates, making some kind of adjustment increases the estimates by 10–15%, with very modest variation due to method at each truncation level. The increase due to the adjustments with 4 Gy truncation is only modestly smaller than that for truncation at 6 Gy kerma. The apparent city difference is larger with the previous adjustment method than for either no adjustment or the various other methods. This is because the Nagasaki adjustment was somewhat inappropriate under the previous method. The dose–response curvature is the ratio of the quadratic to linear coefficients in an LQ model fitted to the dose range 0–1.5 Sv, this restriction being made to stay away from the plateau in the dose response. As expected, the adjustments increase the apparent curvature, but the choice of method has very little effect.

ADDITIONAL EFFECTS OF DOSE UNCERTAINTY

Although substituting $E(x|z)$ for z in the analysis largely removes the bias in the risk estimates, there is some resulting ''overdispersion'' introduced by the random term $x - E(x|z)$. This can affect both the methods for and efficiency of estimating radiation risk, compared to those suitable if true doses were known, and more importantly the standard errors of risk estimates. However, as indicated in refs. (*1–3*), and again below, for the analysis of response times involved in the cancer risk estimation, this overdispersion is negligible, even when the assumed magnitude of errors is increased by introduction of the averaging type. It is important, though, that for some other types of analyses at RERF, particularly those involving chromosome aberration data, this overdispersion will not be negligible, and some allowance for it is required.

Consider, with some simplification regarding time at risk, the relevant data for an individual in terms of a model for the binary indicator of cancer

$$\begin{aligned} y &= \alpha + \beta x + e \\ &= \alpha + \beta[E(x|z)] + \beta[x - E(x|z)] + e \\ &= \alpha + \beta[E(x|z)] + u + e, \end{aligned}$$

where e and u are statistical error terms, with $u = \beta\{x - E(x|z)\}$ representing the overdispersion term referred to above. The point is that because the data on an individual provides quite limited information regarding cancer risk, we find that var(e) is much larger than var(u). Approximating the distribution of y as Poisson and writing the ERR as $\rho = \beta/\alpha$, we have that $\text{var}(u)/\text{var}(e) = \beta^2\text{var}(x|z)/\{\alpha + \beta E(x|z)\} = \alpha\rho^2\ \text{var}(x|z)/\{1 + \rho E(x|z)\}$, which for the RERF data on all solid cancers together is less than about 0.02 for all values of z.

Consequences of this are that to an adequate approximation, (a) statistical methods that would be appropriate if x were known remain appropriate with the substitution of $E(x|z)$, (b) the standard errors from the usual analysis in this sense are appropriate, and (c) to the extent that $E(x|z)$ can be assessed, little information about β is lost from not knowing the true doses.

The essential reason for all of this is that binary observations basically have a large coefficient of variation, which dominates the additional variation contributed by the random term $x - E(x|z)$. There are, however, other types of data used at RERF where this does not apply, and special

procedures to allow for the overdispersion are important. One setting where this arises involves analysis of chromosome aberrations, where the data on an individual are not binary but are the number of aberrant cells (or the number of aberrations) in a sample of about 100 cells from a blood specimen. In this case the model above can be more usefully considered as

$$\begin{aligned} y/m &= \alpha + \beta x + e \\ &= \alpha + \beta[E(x|z)] + \beta[x - E(x|z)] + e \\ &= \alpha + \beta[E(x|z)] + u + e, \end{aligned}$$

where y is the number of aberrant cells within m cells (about 100) examined for each person, and corresponding to the binomial distribution, $SD(e) = [p(1 - p)/m]^{1/2}$ with $p = \alpha + \beta E(x|z)$. Now the standard deviation of u does not depend on m, whereas the standard deviation of e is proportional to $m^{-1/2}$, and for sufficiently large m the overdispersion is no longer negligible.

As indicated in the previous section, the methods developed in ref. (*3*) provide for assessment of $E(x^2|z)$, and hence $\mathrm{var}(x|z)$, in terms of the standard deviations σ_M and σ_A. From this it was found that the approximation $E(x^2|z) = 1.15\{E(x|z)\}^2$ is excellent for the RERF data, from which it follows that $\mathrm{var}(u) = 0.15\beta^2 p^2$. Iteratively reweighted least squares then can be used for fitting the model above and also for assessment of the standard error of risk estimates.

What has been done for some years for this setting is related to this but is preferable in being less model-dependent. Binomial overdispersion in the chromosome aberration data, for whatever reason, is quite apparent and some years ago exploratory data analysis was done to assess empirically the form of $SD(u)$. At that time there had been little investigation of detailed effects of dosimetry errors, and the insight regarding $SD(u)$ provided above was not available. The model that was arrived at empirically, used e.g. in ref. (*10*), had the form $SD(u) = \theta p$, where the parameter θ was to be estimated from the data under analysis, and this can be seen to be compatible with indications given above regarding the form of overdispersion.

RECOMMENDED VALUES OF PARAMETERS

We now turn to the matter of what values of parameters (σ_M, σ_A) should be used. In the original paper (*1*) developing the rationale for allowing for dose errors, it was suggested that previous considerations at RERF indicated that it would be inappropriate to consider errors with coefficient of variation as small as 30%, and it was recommended that ''for the time being'' there should be focus on a value of 35%. Since then there have been several developments suggesting that a value of 35% might be rather small.

The binomial overdispersion in chromosome data referred to above is certainly due in part to dose estimation errors but may also be partly due to variation in radiation sensitivity among survivors. Dose estimation errors with coefficient of variation 50% would explain a substantial part but not all of this overdispersion (*10*). Further, the chromosome aberration dose response is steeper for those who reported severe epilation than for those who did not. Again, this phenomenon is partly due to dose estimation errors, since among those at the same estimated dose those with epilation had higher true dose than others. However, the phenomenon may be partly due to variations in sensitivity to radiation. In ref. (*11*) it was calculated that assuming 50% coefficient of variation for dose estimation errors would be required to explain most of the phenomenon seen in regard to epilation. Similar results, indicating the same level of dose estimation errors, were obtained (*12*) regarding a steeper leukemia dose response for those reporting severe epilation.

Generally, these considerations lead us to conclude that it will be suitable to turn to an error model with larger than 35% coefficient of variation. Moreover, it was not possible in earlier work to consider both measurement and averaging-type errors as dealt with above. It is certainly true that both components of error are present. However, we can see from current results indicated in Fig. 1 that to assume 35% level measurement errors and in addition some level up to 35% averaging-type errors will result in smaller adjustments than those under purely 35% measurement errors. Results discussed above indicate that it does not seem advisable to move to a scheme with even smaller adjustments than have been used until now.

Reference (*4*), available on the website www.rerf.or.jp under Library, List of Publications, provides an analysis of the magnitude of dose estimation errors and was prepared by some of those instrumental in developing and implementing the DS86 system. They argue that to a large extent the revision from the previous DS86 is not expected to reduce the sort of non-systematic, i.e. random, errors we are considering by much. (Incidentally, they consider systematic errors not dealt with here, suggesting that they contribute a coefficient of variation of about 12%.) Their assessment tentatively arrives at a coefficient of variation of about 25% random errors, but all things considered we do not see this as being in serious conflict with the somewhat larger level we will recommend below. In particular, they say on p. 983 that ''there is insufficient information on which to base development of uncertainty values pertaining to errors in input information'', which is a primary component of what we call measurement error. On p. 992 are shown some relationships between DS02 estimates and those based on the tooth enamel activations and chromosome aberrations. Some further analysis of the data underlying these plots would be quite useful, but what is shown is reasonably compatible with our recommendations.

In view of all this, along with results in the previous section, we can with reasonable confidence recommend for the present parameter values of $\sigma_M = 0.40$ and $\sigma_A = 0.20$. The resulting overall coefficient of variation would be $\sqrt{0.40^2 + 0.20^2} = 0.44$. We arrived at this decomposition partly on the grounds that measurement errors are largely unavoidable and difficult to assess, whereas averaging errors are largely introduced intentionally to simplify the do-

simetry system and could have been reduced if deemed too large. The uncertainty estimates in ref. (*4*) pertain in large part to averaging errors, so their assessment tends to support this level of averaging errors.

DISCUSSION

RERF statisticians and radiation scientists involved in biological dosimetry have felt for some time that assuming 35% level dose estimation errors may err somewhat on the low side. It is helpful now to have more explicit means to explore such assumptions, in that the required adjustment factors given by Eqs. (1) and (2) involve explicitly and simply the assumed coefficients of variation for dose estimation, in marked contrast to methods used before. Further, it is helpful to now have a means to consider both measurement and averaging-type errors, and the theoretical results regarding this have been rather surprising. It is well known that averaging errors alone would cause little downward bias in risk estimation, but it was not realized that adding a component of this type of errors can actually reduce the bias. The modeling in ref. (*3*) leading to this conclusion is rather idealized, but once this is suggested by the mathematics we can find an intuitive explanation. This involves the fact that whereas measurement errors cause the estimated doses to be more variable than the true doses, averaging-type errors cause the opposite effect, tending to reduce some of the effect of measurement errors.

Since adjustments depend on the apparent distribution of true doses in the cohort, the question arises of whether in analyses of major subcohorts, or related data, one should use the relevant dose distributions for these. Primary subcohorts include the clinical Adult Health Study, and those on whom chromosome aberrations have been assessed. Related data include that on the offspring of survivors. As for the latter, the radiation doses of the parents are used, and no change to methods here is warranted. Regarding subcohorts, it should be noted that if these were selected on the basis of estimated doses z, then no changes are required since $E(x|z)$ is not affected by such selection. In fact, it would then be wrong to use specific results in this paper with the subcohort dose distribution, since the selection affects not only the distribution of true doses but also the model for measurement-type errors. When the primary subcohorts were selected, no dose estimates were available and the selection was made largely on estimated survivor distance from the bombs. Such selection is closer to one based on estimated than on true doses. To an extent the cohort selections were based on reported acute symptoms, and indeed allowing for this would involve matters not considered here. To do this would involve issues raised in paragraph 2 of the previous section regarding steeper dose response for those with acute symptoms. We recommend that the type of adjustments considered here be used for all purposes at RERF.

Allowing for dose errors involves the parameters (σ_M, σ_A) representing the magnitude of the errors, and unfortunately there is little direct information about the true values of these. Thus it is comforting to see in Table 1 that for one of the most primary needs in risk estimation, there is very little variation due to the assumed values we have considered for these parameters. This does not mean that we should avoid efforts to make the most sensible choice but that in some respects the results are comfortably insensitive to this. It is important on general grounds to do the best we can to arrive at a suitable error model and to continue the considerations of this in the future. The RERF data are used in many ways, and it is likely that for some uses the sensitivity to the error model will be somewhat greater than in risk estimation for all solid cancers taken together, as analyzed here. However, some recommendation must be made for use at least until further progress is made, and this was given at the end of the previous section.

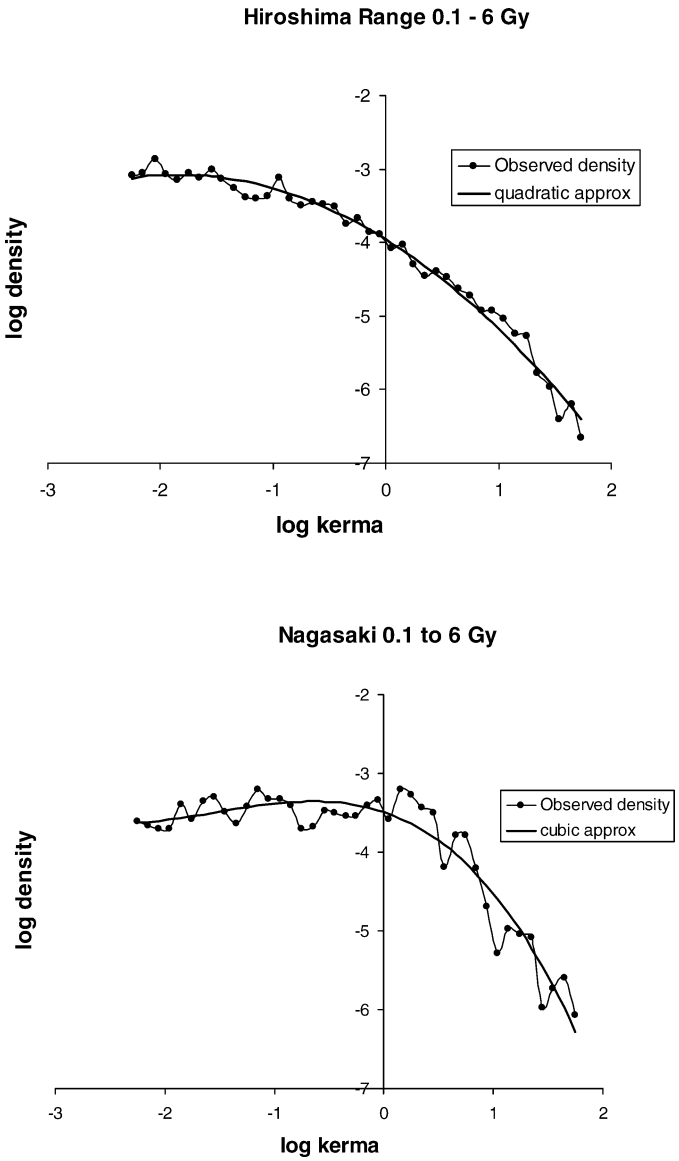

FIG. A1. With minor grouping allowing calculation of the empirical density, graphs for each city of log density as a function of log kerma. For Hiroshima and Nagasaki the curves are respectively quadratic and cubic in log kerma.

APPENDIX

Here we illustrate the assessment of functions $\hat{d}_j(z)$, $j = 1, 2$ for Hiroshima and Nagasaki. First, we fitted smooth curves, suitable for analytical differentiation, to the empirical density functions, represented on a log-log scale, of estimated-dose distributions for each city. Restriction was made to the estimated dose range 0.1 to 6 Gy. This smoothing was done in an exploratory fashion, arriving at results that are quadratic for Hiroshima and cubic for Nagasaki. Results of this are shown in the two panels of Fig. A1. The fits are remarkably good and were essentially the same for DS02 and DS86.

The resulting formulas used in this paper, writing $z^* = \log_e(z)$, are:

Hiroshima: $\text{constant} - 0.956z^* - 0.260(z^*)^2$

Nagasaki: $\text{constant} - 0.471z^* - 0.464(z^*)^2 - 0.1010(z^*)^3$

ACKNOWLEDGMENTS

This publication is based on research performed at the Radiation Effects Research Foundation (RERF), Hiroshima and Nagasaki, Japan, the Oregon Health Sciences University, U.S.A. and the University of Aarhus, Denmark. RERF is a private nonprofit foundation funded equally by the Japanese Ministry of Health, Labour and Welfare, and the U.S. Department of Energy through the National Academy of Sciences. The work at Oregon was partly supported by grant from the National Institute of Health supporting the Oregon Clinical and Translational Research Institute. The work at Aarhus University was partly supported by grant E-2004-AUFF-FLS-1-21 from the Aarhus University Research Foundation.

Received: April 11, 2007; accepted: March 4, 2008

REFERENCES

1. D. A. Pierce, D. O. Stram and M. Vaeth, Allowing for random errors in radiation dose estimates for the atomic bomb survivor data. *Radiat. Res.* **123**, 275–284 (1990).
2. D. A. Pierce, D. O. Stram, M. Vaeth and D. W. Schafer, The errors-in-variables problem: Considerations provided by radiation dose–response analysis of the A-bomb survivor data. *J. Am. Stat. Assoc.* **87**, 351–359 (1992).
3. D. A. Pierce and A. M. Kellerer, Adjusting for covariate errors with nonparametric assessment of the true covariate distribution. *Biometrika* **91**, 863–876 (2004).
4. R. W. Young and G. W. Kerr, Eds., *Reassessment of the Atomic Bomb Radiation Dosimetry for Hiroshima and Nagasaki: Dosimetry System 2002*. Radiation Effects Research Foundation, Hiroshima, 2005.
5. R. J. Carroll, D. Ruppert, L. A. Stefanski and C. Crainiceanu, *Measurement Error in Nonlinear Models: A Modern Perspective,* 2nd ed. Chapman and Hall, London, 2006.
6. G. K. Reeves, D. R. Cox, S. C. Darby and E. Whitley, Some aspects of measurement error in explanatory variables for continuous and binary regression models. *Stat. Med.* **17**, 2157–2177 (1998).
7. D. W. Schafer and E. S. Gilbert, Some statistical implications of dose uncertainty in radiation dose–response analyses. *Radiat. Res.* **166**, 303–312 (2006).
8. D. L. Preston, D. A. Pierce, Y. Shimizu, H. M. Cullings, S. Fujita, S. Funamoto and K. Kodama, Effect of recent changes in atomic bomb survivor dosimetry on cancer mortality risk estimates. *Radiat. Res.* **162**, 377–389 (2004).
9. D. L. Preston, Y. Shimizu, D. A. Pierce, A. Suyama and K. Mabuchi, Studies of mortality of atomic bomb survivors. Report 13: Solid cancer and noncancer disease mortality: 1950–1997. *Radiat. Res.* **160**, 381–407 (2003).
10. Y. Kodama, D. Pawel, N. Nakamura, D. Preston, T. Honda, M. Ito, M. Nakano, S. Funamoto and A. A. Awa, Stable chromosome aberrations in atomic bomb survivors: Results from 25 years of investigation. *Radiat. Res.* **156**, 337–346 (2001).
11. R. Sposto, D. O Stram and A. A. Awa, An estimate of the magnitude of random errors in the DS86 dosimetry from data on chromosome aberrations and severe epilation. *Radiat. Res.* **128**, 157–169 (1991).
12. K. Neriishi, D. O. Stram, M. Vaeth, S. Mizuno and S. Akiba, The observed relationship between the occurrence of acute effects and leukemia mortality among A-bomb survivors. *Radiat. Res.* **125**, 206–213 (1991).