

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

**Donald A. Pierce, Ph.D.; Daniel O. Stram, Ph.D.;
Michael Vaeth, Ph.D.**

RERF Technical Report Series

Technical reports are the basic medium for reporting of original research carried out at the Radiation Effects Research Foundation. Reports in this series receive both internal and external peer review and may serve as the basis for publication in the open scientific literature, in part or in toto. Although they may be quoted and cited, these reports are considered to be internal publications of the Foundation. Copies are available upon request from: Editorial Office, RERF, 5-2 Hijiyama Park, Minami-ku, Hiroshima, 732 Japan.

Beginning in 1989, the RERF Technical Report Series is no longer being published in the traditional Japanese-English bilingual format. However, major reports continue to be available in both languages as separate publications. Selected reports of a highly specialized nature, for which there is presumably less general interest, are produced only in English with an extended Japanese summary.

In this way, the Foundation will be able to more expeditiously report recent findings on the late biological effects of exposure of man to ionizing radiation resulting from the atomic bombings of Hiroshima and Nagasaki.

1989年から、放射線影響研究所の業績報告書は、従来の日英両文を併記した方式では発行しない。主要な報告書については、今後も日英両文で発表するが、それぞれ別に発行する。内容が高度に専門的であり、一般の関心が少ないと思われる報告書については英文のみとし、日本語の要約を添付する。

これにより、広島・長崎の原爆電離放射線被曝の人体に及ぼす晩発性生物学的影響に関する最近の知見を今までよりも速やかにお知らせできることと思う。

The Radiation Effects Research Foundation (formerly ABCC) was established in April 1975 as a private nonprofit Japanese Foundation, supported equally by the Government of Japan through the Ministry of Health and Welfare, and the Government of the United States through the National Academy of Sciences under contract with the Department of Energy.

原爆被爆者データの放射線被曝線量推定値における 確率的誤差に対する補正について[§]

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

Donald A. Pierce, Daniel O. Stram, Michael Vaeth

統計部

要 約

原爆被爆者の放射線被曝個人線量推定値に非系統的誤差が存在するため、線量反応解析において放射線の影響が過小評価され、線量反応推定曲線を歪める結果となっている。線量推定誤差に対して有効な統計モデルがあるという前提で、線形及び二次線量反応モデルの両方について、このような偏りを修正する統計学的方法をここに示す。このような誤差の性質と大きさについては、現在のところ期待されるほど分かっていないのでこの前提は重要である。ここでは、どちらかといえば微妙な統計学上の問題を明確にすることに重点を置いた。この方法は線量推計値の下方修正を含むが、このことはこれらの推定値に偏りがあることを意味するのではなく、むしろリスク推定における偏りを除くための線量反応解析の一部をなすものである。本報告書は主要な焦点を線形線量反応モデルにおいているが、線形・二次モデルの場合の方法も示す。被曝線量推定上の誤差について、信頼できそうな統計モデルを幾つか検討し、誤差の大きさの実質的な範囲を示し、得られた偏りの修正値について感度分析をした。これらの誤差モデルでは、癌死亡の線形過剰リスクの不偏推定値は、誤差を補正しない推定値の約5%~15%増であることが分かった。通常行われているように、極端に高い線量をもつ若干名の被爆者をこの解析から除くと、この幅の上限は約10%減少する。この範囲は極度に広くはなく、また難しさはあるが統計的問題はこの方法を進めるのに十分明確であるので、ほとんどの解析にここに提案する方法による補正を行うことを勧告する。そのような補正の暫定的なものとして、検討した方法のうち中心部分の範囲における特定の誤差モデルを一つ提案する。この勧告にはDS86新線量推定方式に特有の情報は用いられていないが、早急にDS86に関連する情報を組み入れることは重要であろう。被曝線量体系の改正は、主として

[§]全文の日本語訳は別に発行する。

被爆生存者の多数の集団に影響を及ぼす系統的な誤りを修正することを意図しているが、ここに提案した方法は、個々の被爆者の被爆位置と遮蔽状況に関する不完全な情報、つまり、ほとんどが新旧両線量体系に共通な入力情報からもたらされる非系統的な誤差にのみ関与するものである。新線量推定体系の精度に関する今後の情報が極めて有用であるためには、これに関連した微妙な統計学的諸問題に注目することが極めて重要である。特に、真の被曝線量値に属する者についての推定値分布と、ある推定値を与えた者についての真の被曝線量値分布を区別することが肝要である。

Allowing for Random Errors in Radiation Exposure Estimates for the Atomic Bomb Survivor Data[§]

**Donald A. Pierce, Ph.D.; Daniel O. Stram, Ph.D.;
Michael Vaeth, Ph.D.**

Department of Statistics

Summary

The presence of nonsystematic errors in the individual radiation exposure estimates for the atomic bomb survivors results in underestimation of radiation effects in dose-response analyses and also distorts estimates of the shape of the dose-response curve. Statistical methods are presented which will adjust for these biases for linear and quadratic dose-response models, provided that a valid statistical model for the exposure estimation errors is available. This latter qualification is important; less than would be desired is currently known about the nature and magnitude of the errors. Emphasis is placed on clarifying the rather subtle statistical issues involved. The methods involve downward adjustment of dosimetry system estimates, but this is not to imply that these estimates are biased; rather this is part of the dose-response analysis to remove biases in the risk estimates. Primary focus in this report is on linear dose-response models, but methods for linear-quadratic models are also indicated. Some plausible models for the exposure estimation errors are considered, with a substantial range in magnitude of errors, and sensitivity analysis of the resulting bias corrections is provided. It is found that unbiased estimates of linear excess risk for cancer mortality, for these error models, are about 5%-15% greater than estimates making no allowance for such errors. The upper end of this range is reduced to about 10% if, as is commonly done, some survivors with extremely high exposure estimates are eliminated from the analysis. Since the range is not extremely wide and since the statistical issues—although difficult—seem clear enough to proceed, it is recommended that some adjustment of the type proposed here now be made in most analyses. A specific error model in the central range of those considered is tentatively proposed for such adjustments. No information specific to the new DS86 dosimetry is used in this recommendation, even though it will be important to bring such information to bear on DS86 as soon as possible. To a large extent the revision in dosimetry was intended to correct systematic errors affecting large groups of survivors, whereas the methods here pertain only

[§]Full Japanese text will be available separately.

to nonsystematic errors resulting largely from inadequacies in information about location and shielding of individual survivors—input common to old and new dosimetry systems. For further accrual of information regarding precision of the new dosimetry system to be most useful, it is critical that attention be given to the rather subtle statistical issues involved. In particular, it is critical to distinguish between the distribution of estimates for those at a given true exposure, and the distribution of true exposures for those having a given estimate.

Introduction

Much of the research at RERF involves estimation of radiation dose-response for various endpoints among the A-bomb survivors, using estimates of individual radiation exposures which are subject to substantial uncertainty. Focus here is on effects of what might be called “random” errors in exposure estimates, due largely to uncertainties in the survivors’ location and shielding, rather than those of a more systematic nature, such as in the yields of the bombs. It is well known, both from general statistical principles and from specific investigations for this setting, that these random errors do not “average out” in dose-response analyses. Rather they cause systematic biases, both in linear risk estimates and in estimation of the shape of the dose-response. The aim here is to present statistical methods for reducing these biases. For simplicity, most of the development here will be in terms of models linear in true dose, resulting in methods to reduce the bias of the slope parameter. However, it is not difficult to extend the approach to estimation of the shape of the dose-response, in the context of models also containing a quadratic component, and methods for this will be indicated. Further details on this are given in Pierce and Vaeth.¹

The errors-in-doses problem has been investigated in some detail by Jablon² and Gilbert.³ The former report is particularly useful for the analysis given of the probable form and magnitude of errors in exposure estimates. The conclusions of the latter paper are quite similar to those of this report, but the details of development here and the proposed implementation of statistical methods are substantially different. Prentice^{4,5} has discussed special issues arising in this regard in analysis of survival data, which is of primary interest in studying increased cancer risks.

A major purpose here is to lay out the rationale and approach to the problem, which involves both subtle general issues that are not prominent in the mainstream statistical literature and other aspects more special to these data. Some more concrete results are also given but they depend, of course, on specific assumptions regarding statistical models for the errors. No attempt is made here to investigate the actual nature of the errors, but some general comments on this are required.

First, detailed consideration of the nature of the errors is very complex; including the consideration referred to above of systematic versus nonsystematic errors. Statistical models for errors, such as those used here, are certainly an oversimplification, but the other extreme of emphasizing details and complexity

may pose even greater limitations. Further, there has recently been a major revision in the dosimetry system.⁶ The primary change in the new DS86 system is to correct quite systematic errors in the previous system, but there are differences which may affect the structure of nonsystematic errors. Some preliminary discussion of imprecision in the new estimates, as well as a useful summary of the entire system, is given in Chapter 9 of the DS86 report.⁶ It will be important to have further information regarding the precision of the DS86 estimates, including not only general levels, but specifically for shielding categories or even more detailed strata. The implementation of the approach suggested here should certainly evolve as more information becomes available. Research on both refinements of the statistical methods suggested here, and on ascertainment of the actual nature and magnitude of errors in the DS86 estimates, should be considered of high priority at RERF and elsewhere. At the present time, however, it does not seem unreasonable to continue to rely substantially on the assessment of errors given by Jablon² for the previous dosimetry system. He tried to quantify the magnitude of the uncertainty in the input parameters (location of survivor, type of shielding, etc.) for the dosimetry system and evaluated how such random errors propagated to the uncertainty in the dose estimates. The main arguments are also valid for the new dosimetry system, thus the type of model Jablon suggested is taken as a starting point here for sensitivity analysis to assumptions about error models.

In general statistical theory for relating a response variable y to an explanatory variable x , the usual focus is on random variation, or "errors", in the y variable. Situations where x is also measured with error are the subject of a more specialized, but nevertheless vast, literature. Although this certainly provides useful background information, there are special features of the A-bomb survivor data which make this literature less directly useful than might be thought. A generally useful classical reference is Cochran,⁷ and an entry into the most recent literature is provided by papers from a recent workshop.⁸ A recent textbook with broad coverage has been written by Fuller.⁹ Methods particularly related to those discussed here are investigated by Armstrong,¹⁰ Schafer,¹¹ and Prentice^{4,5,12} and reviewed by Clayton.¹³ The latter paper presents a unified approach to statistical modeling of data from epidemiological studies with inaccurately measured exposures. Progress, especially on the computational methods for the analysis of this general model, is still needed. Further developments of these aspects may suggest improvements to the methodology outlined in this report.

The essence of the proposed statistical procedures involves adjustment of the exposure estimates before using them in (linear) dose-response analyses. This is preferable, on a number of grounds, to carrying out the dose-response analysis with the original estimates and afterward adjusting the parameter estimates. Except for low exposures, these adjustments nonlinearly decrease the original estimates to an extent which increases, relative to the magnitude of the estimate, with estimated exposure. Such a procedure is naturally somewhat controversial,

in view of the care which has gone into design of the DS86 system. Because of this, the rather subtle basic ideas involved have been indicated at the outset, before proceeding with the more technical details of the development.

Three of the most basic ideas involved are:

- i) the average in the cohort of the *estimated exposures*, z , among those having approximately the same *true exposure*, x , written as $\text{Avg}(z:x)$,
- ii) the average in the cohort of the *true exposures*, among those having approximately the same *estimated exposure*, written as $\text{Avg}(x:z)$, and
- iii) the average values, or the bias, of parameter estimates for linear models, averaging with respect to both basic sampling variation and random errors in exposure estimates.

Regarding i), a reasonable aim of a dosimetry system is something in principle like $\text{Avg}(z:x) = x$, for all values of x . The primary reasons for saying "in principle" are that: a) only some reasonable approximation to this is an essential issue, and b) it might be considered quite acceptable if this holds true when averaging on a logarithmic (or some other) scale, in which case it will not hold true on the original scale. A useful way to think of the general issues involved in point i) is in terms of an idealized dosimetry system which would give exact estimates aside from errors propagated from those of input parameters such as the individual's location, shielding, etc. From this view, appropriate estimation of these input parameters would lead, at least to a reasonable approximation particularly on a log scale, to $\text{Avg}(z:x) = x$.

However, it is a basic statistical result, discussed in all the references above, that this will not ordinarily imply unbiasedness of parameter estimates in dose-response estimation, which is the issue in point iii). It might be said that the bias in linear risk estimation based on estimated exposure z is directly related to the extent that $\text{Avg}(x:z)$, the average referred to in ii) above, differs from z . It will be shown here that there is certainly a "bias" in this regard, in that $\text{Avg}(x:z)$ is substantially less than z in the higher range of estimated exposure. Moreover, it will be seen that correcting for this form of bias is the key to eliminating the bias of primary concern, that of the parameter estimates.

That $\text{Avg}(x:z)$ is systematically less than z does not mean that the dosimetry system is "biased" in a pejorative sense; indeed it would be quite unreasonable to attempt to devise a dosimetry system which avoids that phenomenon. In particular, for this setting at least, unbiasedness in the sense of i), which should be a primary aim, and that in the sense of ii) and iii), are virtually incompatible. For the A-bomb survivor data, the number of individuals drops off very sharply with increasing exposure. This fact causes $\text{Avg}(x:z)$ to be considerably smaller than z in the upper range of estimated exposure. It is unfortunate that the term "bias" has strong overtones in everyday use, as well as technical meanings which involve some subtlety.

The statistical methods required to deal with the problem involve the replacement, in dose-response analyses, of the estimated exposures z by adjusted estimates $\text{Avg}(x:z)$. This approach is valid for linear (relative) risk models. It is hoped that the discussion just given will make clear that this is not intended to correct any "flaw" in the dosimetry system. Perhaps the simplest way to avoid confusion in this is to think of the adjustments to the exposure estimates as part of the dose-response analyses, rather than as part of the dosimetry system. This viewpoint is also supported by the fact that there are further modifications to the dose-response analyses which must be considered, in dealing with the errors in exposure estimates. Such modifications include an additional term in the variance of the response when analyzing binomial dose-response models. Moreover, the simple solution, replacement of z by $\text{Avg}(x:z)$, is not valid for nonlinear models.

The following artificial example illustrates in a qualitative way some of these basic ideas. The numbers of survivors in the dose categories here are roughly those of the RERF Life Span Study (LSS) cohort, but the model for errors is not intended to be realistic. Rounding true exposures and estimates to the nearest gray, suppose that the statistical model for errors was that with probability 0.5 the estimate was correct, with a probability of 0.25 of having a 1 Gy error in either direction. Table 1 indicates the expected numbers of survivors in a cross-classification of true and estimated exposures. That the estimates are unbiased in the usual sense of i) relates to consideration of the *rows* of this table. On the other hand, $\text{Avg}(x:z)$ pertains to *column* averages, and some selected values are given at the bottom of the table. The "bias" in these, in the sense of ii), is seen to result from far more survivors being 1 Gy below the estimate than 1 Gy above the estimate.

Table 1. Artificial example indicating basic concepts

		Estimate						Number of Survivors
		1	2	3	4	5	6	
True
	2	250	500	250	.	.	.	1000
	3	.	75	150	75	.	.	300
	4	.	.	33	66	33	.	132
	5	.	.	.	15	30	15	60

$\text{Avg}(x:z)$.	.	2.50	3.62	.	.	.

The error distribution in Table 1 was chosen largely for simplicity, but it does serve to illustrate some important points. It is often thought that a major part of the problem resulting from errors is that: i) they are greater for high exposures,

and ii) they may be rather symmetric on a logarithmic scale, resulting in greater overestimates than underestimates. These two factors are relevant, but it is seen from Table 1, having neither of these features, that the fundamental problem would exist even without them.

The method used here for an estimation of $\text{Avg}(x:z)$ in more realistic models is remarkably close in principle to the calculations of Table 1. Estimates can be made of the distribution of true exposures among the survivors, based on the cohort distribution of estimated exposures and a model for the errors. This corresponds in principle to the right-hand column of Table 1. Then the statistical model for errors can be used to distribute survivors at each x across values of z , corresponding to the center part of the table. Then from the resulting cross-tabulation, one can estimate $\text{Avg}(x:z)$ as was done above.

Materials and Methods

This section contains three parts. The first discusses statistical models for errors and a method for estimating $\text{Avg}(x:z)$, to be used as adjusted exposure estimates. The second part indicates why and how adjusted estimates should be used in dose-response analyses and further details the statistical procedures involved in this. The final part describes the data to be used for some example dose-response analyses, in order to indicate the sensitivity of final results to the choice of a model for errors.

Estimation of $\text{Avg}(x:z)$

The required quantity $\text{Avg}(x:z)$ is the mean of the conditional distribution of x among cohort members having (approximately) the same value of z , written as $f(x:z)$. Other quantities computed from $f(x:z)$, such as $\text{Avg}(x^2:z)$ and the standard deviation $\text{SD}(x:z)$ will also be of interest. The concept of this distribution $f(x:z)$ is very central to the development here, and careful consideration of Table 1 perhaps gives the clearest idea of its meaning. It is noted that the probabilistic element here is based neither on thinking of the true exposures to individuals as random variables nor on thinking of the cohort as a sample from a population, but rather it is induced simply by considering the exposure estimation errors as random. This can be seen most concretely in terms of grouped data. If the errors are randomly assigned, then the true exposures of the subcohort belonging to a given z category are thereby distributed randomly over categories of x .

In spite of the concreteness provided by such grouping, there are important advantages in carrying out the development in terms of ungrouped exposures. The most important of these are simplicity and clarity in presentation and the ability to subsequently apply results to whatever specific grouping might be best for a particular analysis. It is perhaps best, though, to think of the development using exact exposures as an idealization which ordinarily will be applied to give results for grouped data.

The density function $f(x;z)$ for the distribution of true exposures among those at a given estimated exposure can be obtained from the relation

$$f(x;z) \propto f(x) f(z;x) \quad , \quad (1)$$

where $f(x)$ is the distribution of true exposures among the cohort, and $f(z;x)$ is the density function of the assumed model for estimation errors. The proportionality is with respect to x , and the proportionality constant, for each z , is simply that which makes $f(x;z)$ integrate to one. In line with the interpretation of $f(x;z)$ discussed above, we emphasize that $f(x)$ is thought of here as the distribution of the true exposures in the cohort, rather than a hypothetical distribution from which the cohort (exposures) is a sample. Choices of $f(x)$ and $f(z;x)$ are discussed below. In this paper, computation of the required quantity $\text{Avg}(x;z)$ is done by numerical integration, so that there will not be artificial constraints on the choice of models for mathematical convenience. After computing $\text{Avg}(x;z)$ for a suitable selection of z values, it was found that the "reduction factors" $[z - \text{Avg}(x;z)]/z$ can be fitted very well by a simple empirical formula, for the error models considered here. These formulas, for any desired error model, can then be used in dose-response analyses to adjust the exposure estimates. In practice, this will often be done to the mean exposures for cells in a cross-classification of the data.

Thus, given an assumed model for estimation errors, $\text{Avg}(x;z)$ can be directly computed if one knew $f(x)$, the distribution of true exposures in the cohort. Because $f(x)$ differs between cities,² the calculation of $\text{Avg}(x;z)$ is done in a city-specific manner. It would not be a very poor approximation to simply take the city-specific distributions $f(x)$ as the distributions of estimated doses, z . It is not difficult, however, to make some adjustment to these for the effect of the errors under an assumed model, and arrive at better estimates of $f(x)$. The approach to this adopted in the present investigation is outlined in the Results section.

The distributions of estimated doses for each city are given in Table 2. The units there are not, strictly speaking, "dose" but DS86 kerma: the tissue kerma free in air at the location of the survivor, adjusted for shielding by structures and terrain, based on the recent DS86 dosimetry.⁶ All the development in this paper is in terms of kerma. Analyses at RERF are sometimes done using organ doses, allowing for shielding by the body. For this case, the recommendation is to compute the above reduction factors in terms of kerma, and then apply these factors to the organ dose estimates. The primary reason for this is simplicity; more direct adjustments of organ doses would require calculations as discussed here for each organ. It is noted that the random errors in the organ doses are somewhat greater than those for kerma, see Chapter 9 of the DS86 report.⁶

The cohort contains a large number of persons at essentially zero exposure, used as a comparison group. The immediate concern here is only with those having nonzero true exposures, and the comparison group can be ignored by making the approximation that the true exposure is zero if and only if the estimated exposure is zero. Further, it has been common practice to truncate the

Table 2. Numbers of survivors whose exposure exceeds selected values of kerma (Gy)

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

exposure estimates at 6 Gy, since survival of a true exposure higher than this is very unlikely and therefore such estimates are particularly unreliable. For the basic development here, such truncation is ignored, although it is recommended that something to this effect be continued in the dose-response analyses. In relation to the present development, the effect of this is to lessen concern about the models for errors, in relation to extremely large exposure estimates.

Attention is now given to the choice of statistical models $f(z;x)$ for the distribution of estimated exposures at a given true exposure. The focal point is that type of model suggested by Jablon,² with substantial motivation given by consideration of major sources of error, in which $\log(z)$ is taken as normally distributed with mean $\log(x)$ and standard deviation independent of x . One key aspect of such models is that on the original scale the errors are proportional to the level of true exposure; this feature seems very appropriate and should probably be incorporated in any alternative model as well. The standard deviation (SD) of $\log(z)$ is approximately equal, for the range of values of interest here, to the standard deviation of z , as a fraction of x ; i.e., the coefficient of variation $CV(z)$. For example, values of 0.30, 0.35, and 0.40 for $SD[\log(z)]$ correspond to 30.7%, 36.1%, and 41.7% for $CV(z)$. For simplicity, a model such as this with $SD(z)=0.30$ will be referred to as "lognormal, 30% error," and so forth. Note that for these lognormal models $CV(z)$ is exactly the square root of $\exp[SD(z)^2] - 1$; the often used approximation $\exp[SD(z)] - 1$, not requiring the normality assumption, is only accurate for smaller values of $SD(z)$ than used here.

Another feature of such models is that the errors are symmetrically distributed on a logarithmic scale, and hence they are skewed towards overestimation on the original scale. The appropriateness of this is suggested primarily by the fact that it is the logarithms of true exposures that are approximately linear in distance from the bombs. As discussed above, the estimates following this model are only unbiased on a logarithmic scale. However, the bias on the original scale due purely to these logarithmic aspects is relatively unimportant in relation to primary issues under consideration here.

Jablon² concluded that the lognormal 30% errors model was worth particular attention, but that models with somewhat greater errors should also be seriously considered. This model is taken as a starting point here, but with emphasis on sensitivity analysis of the effect of possibly larger errors and different models. Results are given for the lognormal 40% errors model and for two types of model other than lognormal, but which have a CV(z) of approximately 40%. Some final numerical results are also given for the lognormal 35% errors model.

The most serious concern with the lognormal model is not the assumption of symmetry on the log scale, but rather the more specific assumption of normality. The normal distribution is noteworthy, indeed notorious, for having very "light tails"; that is, it allows for very few large deviations in relation to the size of more typical ones. It might be suspected that the results here are sensitive to the assumed chance of extreme errors, so investigating a distribution with "heavier tails" is important. A natural choice, called a contaminated normal distribution, is to model $\log(z)$ as being, with a 85% chance, normally distributed with a standard deviation of 0.30, and with a 15% chance, normally distributed with a standard deviation of 0.75. These parameters were chosen such that $\log(z)$ has a standard deviation of 0.40, as in the second model above, but with a quite different distributional shape.

A fourth model was chosen to examine departures from the assumption of symmetry on the log scale, with the consequent skewness towards overestimation. This model takes z as normally distributed, with mean x and $CV(z)=40\%$, providing a useful comparison to the previous two models. The authors feel that approximate symmetry on a log scale is more natural, and that this fourth model is a rather extreme departure from that; but nevertheless provides useful information about the effect of distributional shape.

Thus, results will be given for four error models:

- A. lognormal, 30% error
- B. lognormal, 40% error
- C. contaminated lognormal, 40% error
- D. normal, 40% error

It is natural to consider whether error models should be specific to shielding categories, e.g., unshielded, in a traditional Japanese house, in a heavy commercial building, and so forth. Allowing the percent error to depend on these factors, and perhaps others, is likely to be the most important extension of the methods developed here. Future developments of the DS86 dosimetry system will provide the information necessary for evaluation if such a refinement of the present methodology is required.

Use of these adjustments for dose-response analyses

First, models in which the expected response is linear in true exposure are considered. The following indicates why replacing exposure estimates z

by adjusted exposure estimates $\text{Avg}(x:z)$, for the types of data and statistical methods of concern, will reduce the bias in parameter estimates. Some comments on nonlinear models are given at the end of this section. The definition of "bias" here involves averaging over both ordinary sampling variation and random errors in exposure estimates. Despite the reduction of bias by this substitution, it is clearly not the case that the quantities $\text{Avg}(x:z)$ should be thought of as the equivalent of the true x 's. Thus, it is important to investigate whether additional allowances in analysis or interpretation are required. An attempt is made to focus only on essential points, with some technical details relegated to an Appendix.

Before proceeding with the discussion, it should be noted that the same adjustments $\text{Avg}(x:z)$ should be used whether an analysis is of the entire cohort or, as is often the case, a subset of it selected on the basis of estimated exposures. This question arises because the distribution $f(x)$ of true exposures in the entire (exposed) cohort is involved in the calculations, and it might be thought that the restriction of this to some subcohort should be used in such cases. That this is not so is because the selection must be made on the basis of the z values, rather than the x values. The distribution $f(x:z)$ does not depend on how z is selected, and thus the development here applies with no changes for analysis of selected individuals.

In the following development, exposure and city will be taken as the only covariables, but the incorporation of additional covariables measured without error does not alter the basic results provided that neither the distribution of true exposures nor the nature of estimation errors are related to such covariables. The primary additional covariables of interest in the LSS are city, sex, and age-at-exposure. The effects of city are accounted for by calculating the adjustments $\text{Avg}(x:z)$ in a city-specific manner, as indicated above. In principle, one should also consider sex and age-at-exposure specific adjustments. However, neither of the model aspects mentioned above differs substantially by sex or age-at-exposure, so the effect of these additional adjustments would be minimal.

Whether further modification in standard methods is required, in addition to substitution of adjusted exposure estimates, turns out to depend on the type of data being analyzed. Consideration is given to two primary data types of interest at RERF: i) cancer incidence or mortality, involving statistical methods of survival analysis; and ii) chromosomal aberration prevalence, involving methods for binomial data. It will be seen that for the first of these no additional modification to standard methods or interpretations is required. For the latter, though, analysis under a binomial model with adjusted doses is not adequate.

The relevant distinction between these involves the relative magnitudes of ordinary sampling variation and the additional variation due to errors in exposure estimates. For situation i), the datum on each individual is essentially binary, whereas for ii) it is the proportion of a large number of examined cells which exhibit an aberration. In the latter case, the ordinary sampling variation, being

inversely proportional to the number of cells examined for each person, is much smaller than in the former.

For the cancer data, there are some additional complexities since the time to cancer incidence or mortality, being a random variable, is also a part of the data on each individual. However, the results of Prentice⁴ indicate that for linear relative risk models all that is required is a substitution of $\text{Avg}(x:z)$ for z when analyzing cancer data using Cox's regression methodology. The statistical models used at RERF for analyzing cancer data are essentially a grouped-time version of these methods, and it may clarify Prentice's result to express it explicitly in terms of the grouped-data setting. Some further details are provided in the Appendix.

For the chromosomal data, overdispersion relative to the binomial model is quite apparent, and modifications to binomial methods have been used for some time in analyses of these data; see, e.g., Preston et al.¹⁴ Provided that these special methods (further described in the Appendix) are used, the only additional modification required is the substitution of adjusted exposure estimates.

In terms of true exposures, the observation for an individual with exposure x follows a model of form

$$y = \alpha + \beta x + \text{error} \quad , \quad (2)$$

where y is the proportion of aberrant cells in the sample of cells from the individual. The standard maximum likelihood methods for these data can be thought of as weighted least squares analysis under this model. The expectation of "error" is zero, which implies unbiased estimation. The variance of "error" depends on x , calling for weighted regression; it also depends on α and β , calling for iterative methods.

For a survivor selected at random from the cohort with estimated exposure z , (2) can be reexpressed as

$$y = \alpha + \beta \text{Avg}(x:z) + \beta [x - \text{Avg}(x:z)] + \text{error} \quad . \quad (3)$$

Note that x is, in this setting, a random variable with distribution $f(x:z)$, so the term $\beta [x - \text{Avg}(x:z)]$ becomes an additional "error term" in the model, with expected value zero and variance $\beta^2 \text{Var}(x:z)$. The last factor here, the variance of x among those with a given z , thus becomes an important quantity, which can be computed for a given error model very similarly to $\text{Avg}(x:z)$.

Thus, the primary result is that the datum for an individual with estimated exposure z can be expressed as

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

where error^* is the combination of the two error terms in (3). That error^* has an expected value of zero when averaging over both sampling variation and errors

in exposure estimation implies that weighted linear regression analysis based on (4) will yield unbiased estimates in the desired sense. Further, it follows from a standard calculation that

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

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

In the case of the cancer data, the first term in (5) vanishes; i.e., the variance of error^* in (4) is the same function of $\text{Avg}(x:z)$ as the variance of the error in (2) as a function of x . Thus, no changes in analysis methods are required beyond substituting $\text{Avg}(x:z)$ for z . For the chromosomal data, the two terms in (5) are of roughly the same order of magnitude. These results are discussed further in the Appendix, where it is also shown that a special method which has been in use for the chromosomal data is suitable for variances of the form of (5), with no further modification beyond substitution of adjusted dose estimates. Finally, justification is given there for use of weighted least squares for (4), even though, in contrast to (2), this is no longer the maximum likelihood method.

All the above results pertain to models in which the expected response is linear in true exposure. Corresponding development for nonlinear models is in general less tractable, but there is a natural extension for the case of linear-quadratic models, i.e., those where $\alpha + \beta x$ in (2) is extended to $\alpha + \beta x + \gamma x^2$. In this case, unbiased estimation would be achieved by extension of the above to include replacement of z^2 in the analysis by $\text{Avg}(x^2:z)$. This last quantity can be calculated similarly to $\text{Avg}(x:z)$; indeed the need for this has already arisen above in terms of $\text{Var}(x:z)$, which is $\text{Avg}(x^2:z) - [\text{Avg}(x:z)]^2$. Particular caution should be exercised in using this approach, since it may be expected that $\text{Avg}(x^2:z)$ will be more sensitive to the choice of error model than is $\text{Avg}(x:z)$. Further details on parameter estimation in linear-quadratic models with errors in exposure estimates are given in recent work by Pierce and Vaeth¹ on the shape of the cancer mortality dose-response curve. Moreover, the above arguments regarding the additional variation in the data due to errors in exposure estimates will require some extension for nonlinear models. Further discussion of these points will be given in subsequent sections.

Data and models used for sensitivity analysis

The most useful way to analyze the sensitivity to the error models used here is to see what effect the adjustments made using them have on parameter estimates in actual dose-response analyses. For this purpose, analyses of cancer mortality data are made using precisely the data and models of recent analyses by Preston and Pierce.¹⁵ Estimates are computed for all cancers except leukemia as a group, and for leukemia alone, using models linear in exposure for the excess relative risk. In these models, the age-specific excess relative risk is taken as constant in

age, but depending on sex and three categories of age-at-exposure. Summary risk estimates are taken, as by Preston and Pierce for comparing dosimetry systems, as the simple average over these six groups of the risks per unit exposure. For nonleukemia, the analysis is in terms of dose to the large intestine, as representative of the dose to organs with substantial shielding by the body; for leukemia, analysis is in terms of dose to the bone marrow. Adjusted organ dose estimates are computed here by applying the reduction factor computed in terms of kerma to the organ dose estimates provided by the dosimetry system.

Results

Consideration is first given to estimation of $f(x)$ for each city. Among survivors with positive estimates, the proportion above any z value can be fitted very well by formulas of the form $\exp(-\theta_1 z^{\theta_2})$, which is the distribution function of the Weibull model. (The values of θ_2 are less than one, corresponding to monotonically decreasing density function.) An adequate approach seems to be to assume the same parametric form for the distribution of x , and then to estimate the parameters θ_1 and θ_2 for that model using the data in Table 2 and allowing for the effect of errors in the estimates there. This has been done fairly simply by choosing (by a direct search) these parameters so that, with a given model for the errors, the induced theoretical distribution (computed numerically) of the z 's agrees very closely with the observed distribution. The dependence of these answers on the specific error models defined above is not great; for simplicity what will be used for all purposes in the following are the values for a fifth "central" model: lognormal, 35% error. The results are:

$$\begin{array}{ll} \text{Hiroshima:} & \theta_1 = 2.84 \quad \theta_2 = 0.50 \\ \text{Nagasaki:} & \theta_1 = 2.33 \quad \theta_2 = 0.50 \end{array}$$

The equal values of θ_2 , for the two cities, is not due to specifying them so; although it is coincidental, the fitted values did very nearly agree to two decimal places and have only been slightly rounded above. The quality of the fit provided by this choice of parameters is not described in detail here, but it was remarkably good.

With this city-specific model for $f(x)$, it is straightforward to compute $\text{Avg}(x:z)$ numerically. Results for each error model at selected z values are given in Table 3.

In the previous section were noted several reasons for interest in $\text{Var}(x:z)$, the variance of x among those at given z ; and in $\text{Avg}(x^2:z)$, which is given by $\text{Var}(x:z) + [\text{Avg}(x:z)]^2$. Numerical results for this are described most simply in terms of the squared coefficient of variation, $\text{Var}(x:z)/[\text{Avg}(x:z)]^2$, values of which are given in Table 4 for the four error models.

Discussion of the results of Tables 3 and 4 will be given in the following section, but it is briefly noted now that model C gives very similar results to model B. The error models C and D will not be pursued further, and subsequent numerical

Table 3. Adjusted exposure estimates for the four error models

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

A: lognormal, 30% error; B: lognormal, 40% error; C: contaminated lognormal, 40% error; and D: normal, 40% error.

Table 4. Squared coefficient of variation for the four error models

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

A: lognormal, 30% error; B: lognormal, 40% error; C: contaminated lognormal, 40% error; and D: normal 40% error.

results will be given only for the lognormal models A and B, with the addition of an intermediate lognormal model with 35% error.

For use in applications, convenient formulas for Avg(x:z) are required. It has been found that second-degree polynomials in log(z) can be fitted remarkably well to the *reduction factors*, $[z - \text{Avg}(x:z)]/z$. Using these approximations to calculate Avg(x:z) from z is accurate to within 0.01 Gy over the range 0.5–6 Gy. It is proposed that no adjustment be made to estimates under 0.5 Gy and that estimates above 6 Gy be reduced to 6 Gy and then adjusted further as though this were the estimated exposure (which is equivalent to making the adjustment

to the actual estimate and then truncating the result to the adjusted value corresponding to 6 Gy). Coefficients for the approximations to the reduction factors are given in Table 5.

Table 5. Coefficients for calculating reduction factors

	Error Model Lognormal Percent Error	Coefficients for Terms:		
		Constant	$\log(z)$	$[\log(z)]^2$
Hiroshima	30%	0.03597	0.05807	0.01166
	35%	0.04732	0.07623	0.01336
	40%	0.06036	0.09684	0.01314
Nagasaki	30%	0.01500	0.05304	0.00885
	35%	0.01900	0.06545	0.01374
	40%	0.02817	0.08031	0.01558

Although the values in Table 4 can also be fitted very well by such polynomials, they are remarkably constant in z and it would be an overinterpretation of the models used here to emphasize the small variation seen for a given error model. That the numbers there are small, in the range 0.09–0.15, is important. The additional variation in the chromosome aberration data due to errors in exposure estimates depends on these values, as indicated above and in more detail in the Appendix. Further, adding 1.0 to the numbers in Table 4 provides the value of $\text{Avg}(x^2:z)/[\text{Avg}(x:z)]^2$, which is relevant to the fitting of models quadratic in true exposure. Finally, the square roots of the numbers in Table 4 are the standard errors, as a fraction of $\text{Avg}(x:z)$, of the distributions of x for a given z value, which are of some general interest in interpretation of exposure estimates.

The standard deviation of a distribution does not give a complete description of variability, unless the distribution is normal or other information about the shape is known. In this regard, graphs are given in Figure 1 of some examples of the distributions $f(x:z)$. Also given on the same figures are corresponding graphs of the distributions $f(z:x)$, in order to emphasize the distinction between the two types of distributions important here. Note that the horizontal axis for each figure corresponds to x for one graph and to z for the other.

Comparison of the values of $\text{Avg}(x:z)$ above for various error models does not give a very clear picture of the sensitivity analysis in practical terms, since the relative importance of parts of the dose range in fitting dose-response models is unclear (and difficult to describe). Thus Table 6 is provided to indicate the effect of using the various error models for some actual linear dose-response analyses. Details of the data, models, and the way of summarizing parameter estimates were given in the previous section. Shown are the relative increases in the same summary parameter estimates used by Preston and Pierce¹⁵ to describe the

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

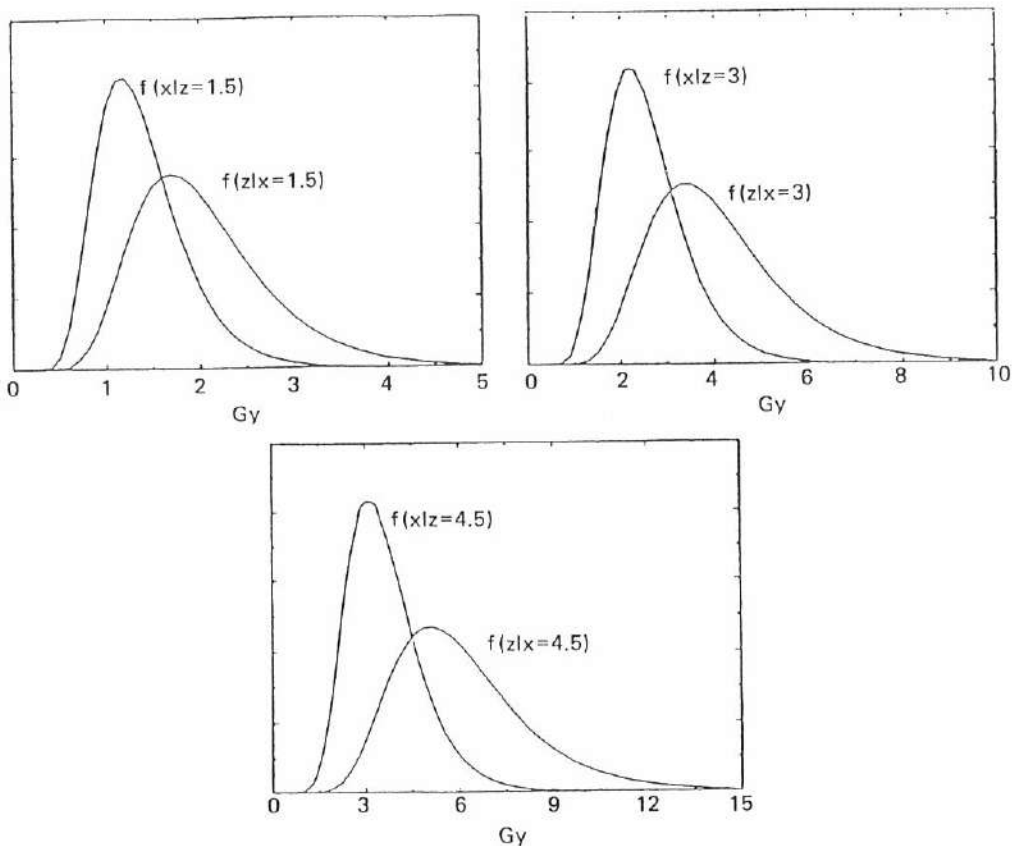


Table 6. Increases in cancer risk estimates for three lognormal error models

	Error Model Lognormal Percent Error	Increase in Estimated Excess Relative Risk	
		0-6 Gy	0-4 Gy
All Cancer Except Leukemia	30%	10.0%	6.8%
	35%	13.3%	9.0%
	40%	16.7%	11.4%
Leukemia	30%	6.1%	4.3%
	35%	8.1%	5.6%
	40%	10.2%	7.2%

effects of the recent revision in dosimetry. As in that paper, two types of estimates are given, the comparison of which is quite relevant for present purposes and is discussed later. These consist of using the entire cohort with kerma estimates limited to 6 Gy, and alternatively, restricting analysis to those with estimates below 4 Gy. Very specifically, the percentage increases in Table 6 apply to the bottom two entries in each of Tables 11 and 12 of Preston and Pierce.¹⁵

For the cancer data, the standard errors of the parameter estimates relative to the estimates, i.e., the coefficients of variation, are essentially unchanged by the analysis allowing for dose-estimation errors. This reflects the fact that the new procedure does not require any change in the model for the variance of the response, but only an adjustment of estimated doses. For the chromosomal data, on the other hand, one would expect that analysis with adjusted doses would lead to larger standard errors relative to those under the binomial model, since here the necessary modifications include adding an overdispersion term to the variance of the response. Methods allowing for overdispersion relative to the binomial model have already been in use, however, and the standard errors under the proposed method will be very similar to those given by current methods.

Discussion

As noted at the outset, the method developed here is not fundamentally different from that given by Gilbert³ but the implementation is more suitable for many purposes, especially since it is not linked to extensive calculations for a particular choice of dose categories. The principles are also consistent with those discussed by Prentice,^{4,5} who deals more specifically with application to survival data, but here much more emphasis is placed on the details specific to the LSS setting. The general approach laid out by Clayton¹³ seeks to establish a unified solution to the problem of inference in the three components of the model: the dose-response model, the errors in exposures, and the distribution of exposures. This may certainly have some theoretical advantages, but the computational methods are still in the developing stages. In the present approach, these three aspects are considered separately, thereby reducing the the computational problems considerably. However, improvements should be considered as the implementation of Clayton's general model progresses. Much of the literature on regression problems with measurement errors in the explanatory variable focuses on methods to adjust the parameter estimates obtained in an analysis using the original exposure estimates. These methods typically require special assumptions about the form of the distribution of true exposures, which are not appropriate to the LSS data. Moreover, any particular study would entail a different parameter adjustment procedure, tailored to the distribution of observed dose in the specific subsample of the LSS cohort being investigated. It seems clear that, at least for models suitably linear in true exposure, the approach taken here leads to an appropriate method for reducing, from parameter estimates in dose-response analyses, the bias resulting from random errors in exposure estimates. That is, these methods would essentially eliminate such biases if the true statistical model for errors in exposure estimates were known; and they can be used to

substantially reduce them by using an approximate model. It is emphasized again that the adjustments required to the exposure estimates do not imply a defect in the dosimetry system and that the methods here are only to deal with nonsystematic errors in exposure estimates.

The only major difficulty, which is of course unavoidable, involves the required assumptions regarding the statistical model for errors in exposure estimates. The remainder of the discussion here will deal primarily with interpreting the sensitivity analysis in regard to such models, and with some issues arising in consideration of nonlinear models.

First, it is emphasized that for purposes of linear dose-response analyses the displays in Tables 3 and 4 may both give a false impression of the extent of bias due to exposure estimation errors and overemphasize the sensitivity to the choice of error model. As may be seen in Table 2, there are relatively few survivors in the upper part of this exposure range. Thus, for many purposes the comparisons in Table 6 provide a more useful basis for conclusions regarding both issues.

Tables 3 and 4 do indicate that there is very little difference between the use of the lognormal model and the contaminated lognormal model, which has substantially "heavier tails". Thus, although the latter may have substantial appeal as a more realistic model, there seems to be little need to introduce the additional complexity. The normal model with constant coefficient of variation does give results somewhat more different, in the sense of smaller adjustments. When graphs of $f(x:z)$ as in Figure 1 are made for this model, it is seen that the right tails of these distributions are very similar to that in Figure 1 but the left tails are noticeably shorter. This reflects an aspect of this model which is quite contrary to what might be expected, the understanding of which may lessen concern about the long right tail (on the dose, not log dose, scale) of the lognormal error model. This is that the final factor in equation (1), $f(z:x)$, thought of as a function of x for fixed z , is even more skewed to the right for the normal model with constant coefficient of variation than it is for the lognormal model. That is, in statistical terms, even though the distribution of z given x is symmetric for the normal model, the "likelihood function" for x given z is quite skewed to the right. Thus there are subtleties in the actual distinction between these two types of models. The authors and others who have considered this problem feel that the lognormal model is somewhat preferable both on general statistical grounds and on those which are more specific to the dosimetry system.

The information provided by Table 4 is quite important in several respects. That the entries in Table 4 depend little on z is important for at least two reasons. First, this implies that the special methods used for some time to allow for overdispersion in the chromosomal data are also well suited for the needs here. Secondly, the constancy in z means that $\text{Avg}(x^2:z)$ is approximately proportional to $[\text{Avg}(x:z)]^2$. This implies that fitting models quadratic in $\text{Avg}(x:z)$ will provide good estimates of the shape of the dose-response curve for models quadratic in

true exposure, although the quadratic coefficient will be biased upward by a factor of one plus the quantity given in Table 4. Generally speaking, though, further analysis along the lines of this report for the case of nonlinear models may be useful. Pierce and Vaeth¹ study the effect of errors in exposure estimates for inferences about linear-quadratic models, with particular focus on the extent to which apparent departures from linearity may be due to such errors.

Table 6 indicates, among other things, the extent to which the effect of errors in exposure estimates is smaller when those with exposures above 4 Gy are eliminated from the analysis. As discussed by Preston and Pierce¹⁵ and Pierce and Vaeth,¹ though, this is not the primary motivation for this restriction in practice. The more critical issues involved are that the true dose-response may not be linear over a wider range and that risk estimation at low exposures is the predominant concern. Nevertheless, from the viewpoint of this paper, such a restriction does serve to lessen concern about the appropriateness of an error model as it relates to extremely high exposure estimates.

There are issues in addition to bias in dose-response analyses which are related to errors in exposure estimates. In particular, there is substantial interest in the extent to which there may be biological variation among individuals in their general sensitivity to irradiation. This is potentially relevant to the cancer risk estimation, since substantial variation of this type might imply that those who survived the acute effects would be less sensitive, in terms of cancer risk, than a general population. One way to investigate this issue is, in essence, to study whether there are any positive associations between different radiation effects (cancer, acute symptoms, chromosomal aberrations, etc.) among those at the same exposure level; see, e.g., Ban et al,¹⁶ Hakoda et al.¹⁷ It is certain, though, that random errors in exposure estimates will induce spurious positive association of this nature. Among those at (approximately) the same estimated exposure, there is in fact a distribution of true exposures, and the risk of the various radiation effects will depend on these true exposures rather than on the estimated ones. Thus, valid investigations along these lines will have to consider whether any positive associations found are greater than those which would result from the errors in exposure estimates. The statistical formulation of this report can be extended to calculations of this nature.

In conclusion, although further research in this area is important, it seems advisable to begin making use of the methodology developed in the present report in most analyses of these data. The most important future improvements will involve refinements on the model for exposure estimation errors, and thus ascertainment by RERF and dosimetry experts of the magnitude of the impression in the DS86 system should be given high priority. Emphases in this should include investigation of the dependence of the magnitude of errors on shielding category and distinction between random and systematic errors. This is a difficult task, however, and the sensitivity analysis given here suggests that while

awaiting this, use of an error model along the lines suggested below will, for many purposes, be much better than making no adjustments at all.

Although there may be other statistical approaches to removing the biases, such as those briefly discussed above, the approach taken here of adjusting exposure estimates seems eminently suited for the particular needs arising with these data. This is primarily because the exposure estimates must be used for a wide variety of purposes: linear risk estimation for different types of cancer, chromosome aberrations, noncancer mortality, clinical endpoints, genetic effects, and acute effects; and for fitting linear-quadratic models to investigate the shape of the dose-response for several of these endpoints.

When feasible, particularly for certain purposes such as analysis of the shape of the dose-response, it will be appropriate to pay special attention to the choice of an error model and to carry out further sensitivity analysis. Many reports, however, are already complicated by attention to a large number of issues, and such special attention to error models will be infeasible. These investigations ordinarily use linear dose-response models, and the sensitivity analyses here suggest that the uncertainty in parameter estimates due to choice of an error model may be relatively small in relation to other uncertainties. For these it may be best to rely primarily on use of a given, simple, choice of an error model, with some brief indication of the effect of the adjustment. It should be recognized that the adjustments for error are an approximation, similar to other approximations made in order to arrive at a useful analysis; e.g., the use of linear risk estimates.

With these thoughts in mind, the authors feel that it will serve a useful purpose to suggest an error model to be used at this time, for analyses where it is infeasible to devote substantial special attention to the issue. The above results and discussion, as well as previous analyses and more basic statistical considerations, suggest that a lognormal model should be used for this purpose. It seems likely to the authors that the choice of 30% errors may be somewhat optimistic. A primary part of the basis for this view is the judgment of Jablon² that "sources identified in this section amount to about 30% plus or minus," preceded by the comment "reflecting on how they [certain aspects] were derived it seems not unlikely that we have underestimated." The uncertainties that Jablon considered were largely those due to inaccuracies in the assessment of the survivors location and shielding, inputs which are common to the old and the new dosimetry systems. Errors in the DS86 system somewhat less than 30% are now being discussed,⁶ but it appears that to some extent this may be in reference to the technical errors arising from the dosimetry system itself, taking the input parameters of survivors location and shielding at face value.

On these grounds, the authors suggest that for the time being some special focus be placed on the lognormal, 35% errors, model. Since numerical values for exposure adjustments under this model were not given in Table 3, although the formula is given in Table 5, these are indicated here.

Estimate	0.50	1.00	2.00	3.00	4.00	5.00	6.00
Adjusted Value							
Hiroshima	0.50	0.95	1.79	2.56	3.28	3.98	4.64
Nagasaki	0.50	0.98	1.86	2.68	3.46	4.20	4.92

The values corresponding to those in Table 4 for this model are in the range 0.11–0.12.

Appendix

Some additional details regarding the dose-response analysis are given here.

For the cancer data, consideration is given only to proportional hazards models with relative risk linear in dose. The binary observation y for an individual with true exposure x can be written as

$$y = S [\alpha + \beta x] + \text{error} \quad (1)$$

where S is the integral of the underlying hazard over the time at risk for the individual, a random variable. The value S may include parameters to be estimated, often relating to stratification. Since the time at risk is random, the mean and the variance of "error" in (1) are defined in terms of the quantity $[y - S (\alpha + \beta x)]$. The mean of this is zero, and the variance is the expected value of S times $[\alpha + \beta x]$. Iterative weighted least squares in (1), taking the variance as slightly modified to be $p_x = S [\alpha + \beta x]$, is equivalent to maximum likelihood estimation in this model. This is referred to as "Poisson regression analysis," since it would also arise if S were thought of as fixed and y was taken to have a Poisson distribution with mean $S [\alpha + \beta x]$.

For the chromosomal aberration data, under a binomial model, the observation for an individual with true exposure x can be written as

$$y = \alpha + \beta x + \text{error} \quad (2)$$

with $\text{Var}(\text{error}) = p_x (1 - p_x)/n$, where $p_x = \alpha + \beta x$ and n is the number of cells examined for each individual. Again, iterative weighted least squares is equivalent to maximum likelihood estimation.

For the case of errors in exposure estimates, the data for a survivor selected at random from the cohort having estimated exposure z can be expressed as follows, for the two cases:

$$y = S [\alpha + \beta \text{Avg}(x:z)] + \text{error}^* \quad (3)$$

$$y = \alpha + \beta \text{Avg}(x:z) + \text{error}^* \quad (4)$$

Here, error* involves both the error in (1) or (2) and a quantity involving the random variable $[x - \text{Avg}(x:z)]$ under the distribution $f(x:z)$. For (3) there is a very small dependence which will be ignored, between the random variable S and the exposure estimation error, due to the dependence of S on time to cancer; see Prentice.^{4,5} With this approximation, the mean of error* is zero for both instances, when averaging over both the sampling variation and the distribution of x given z. Thus, least squares estimation is unbiased in this sense. Straightforward but somewhat lengthy calculations yield that

$$\text{Var}(\text{error}^*) = \begin{cases} p_z & \text{[for (3)]}, \\ p_z(1 - p_z)/n + \beta^2 [(n - 1)/n] \text{Var}(x:z) & \text{[for (4)]}, \end{cases}$$

where, with a slight abuse of previous notation for p_x , we write $p_z = S[\alpha + \beta \text{Avg}(x:z)]$ and $[\alpha + \beta \text{Avg}(x:z)]$ for the two settings, respectively. For the cancer data case, p_z is not strictly speaking the variance of error* but, in the same way as for (1) above, involves replacing the expected value of S by the observed S. The important point is that the weights for regression in (3) are the *same* function of $\text{Avg}(x:z)$ as the function of x used in the standard methods for (1); thus after having replaced z by $\text{Avg}(x:z)$ the usual methods apply. This is rather surprising, implying an absence of additional variation in (3) due to replacing the unknown x by $\text{Avg}(x:z)$. The calculations leading to this depend on the fact that S is random, although the apparent overdispersion arising if S is treated as fixed is much smaller than p_z . These calculations also depend on the approximation that S is independent of x in the distribution $f(x:z)$; however, a detailed investigation has shown that the effect of this approximation is extremely small in the present setting.

For the binomial case, the first term represents the variance which would be assumed if the only modification in the standard method was to replace z by $\text{Avg}(x:z)$; the second term represents overdispersion which also needs to be taken into account. Primarily because of the factor n in the binomial variance $p_z(1 - p_z)/n$, and also to some extent because p_z can be large, the overdispersion term is roughly of the same order of magnitude as the binomial variance and therefore cannot be ignored. Since, as is shown in the text, $\text{CV}^2(x:z)$ is approximately constant in z for a given error model, the overdispersion term there is approximately proportional to $\text{Avg}(x:z)^2$. The methods which have been used for some time at RERF involve allowing for overdispersion using a model

$$\text{Var}(\text{error}) = p_x(1 - p_x)/n + \theta p_x^2 \quad ,$$

where θ is estimated from the data. See, e.g., Preston et al.¹⁴ Since α is very small, this is essentially the same as using the model above, with overdispersion proportional to $\text{Var}(x:z)$. Using the existing method with $\text{Avg}(x:z)$ replacing z will be essentially the same as the analysis with overdispersion suggested above, except that the extent of overdispersion, θ , will be estimated from the data rather than through the formulas above involving $\text{Var}(x:z)$. This is quite suitable, indeed preferable, since for data of this type it is natural to suspect that there may be sources of overdispersion in addition to that due to the exposure errors.

As indicated above, these iterative weighted least squares methods are equivalent to maximum likelihood when exposures without error are used. When there are such errors this is not the case, and maximum likelihood methods involve very heavy computations which are unsuitable for routine use. These involve numerical integration of the likelihood function in terms of x , with respect to the distribution $f(x;z)$. It is of interest, though, to see how they compare to the simpler weighted least squares methods described above. Simulations have been carried out, for both of the above types of data in settings which are quite close to the RERF data, which implement both the maximum likelihood and weighted least squares methods. As would be expected on general theoretical grounds, there is very little difference in the results of the two methods. It seems clear that in general there should be no practical motivation for using maximum likelihood methods rather than the weighted least squares proposed here.

References

1. Pierce DA, Vaeth M: The shape of the cancer mortality dose-response curve for the A-bomb survivor data. (RERF TR in preparation)
2. Jablon S: Atomic bomb radiation dose estimate at ABCC. ABCC TR 23-71
3. Gilbert ES: Some effects of random dose measurement errors on the analyses of atomic bomb survivor data. *Radiat Res* 98:591-605, 1984 (RERF TR 12-82)
4. Prentice RL: Covariate measurement error and parameter estimation in Cox's regression model. *Biometrika* 69:331-42, 1982
5. Prentice RL: Covariate measurement errors in the analysis of cohort and case-control studies. In *Survival Analysis*, eds Crowley J and Johnson R, Institute of Mathematical Statistics Monograph Series, Vol 2, 1982
6. Roesch WC (ed): US-Japan Joint Reassessment of Atomic Bomb Radiation Dosimetry in Hiroshima and Nagasaki. Final Report, Volumes 1 and 2. Radiation Effects Research Foundation, Hiroshima, 1987
7. Cochran WG: Errors of measurement in statistics. *Technometrics* 10:637-66, 1968
8. Proceedings of Workshop on Errors-in-Variables in Epidemiology, National Cancer Institute. (To appear in *Statistics in Medicine*, 1988)
9. Fuller WA: *Measurement Error Models*. New York, John Wiley and Sons, 1987
10. Armstrong B: Measurement error in the generalized linear model. *Comm in Statist* 14:529-44, 1985
11. Schafer DW: Covariate measurement errors in generalized linear models. *Biometrika* 74:385-9, 1987
12. Prentice RL: Binary regression using an extended beta-binomial distribution, with discussion of correlation induced by covariate measurement errors. *J Am Statist Assn* 81:321-7, 1986

13. Clayton DG: Models for the analysis of cohort and case-control studies with inaccurately measured exposures. In *Statistical Models for Longitudinal Studies of Health*, ed Dwyer JH. Oxford University Press, 1988
14. Preston DL, McConney ME, Awa AA, Ohtaki K, Itoh M, Honda T: Comparison of dose-response relationships for chromosome aberration frequencies between the T65D and DS86 dosimetries. RERF TR 7-88
15. Preston DL, Pierce DA: The effect of changes in dosimetry on cancer mortality risks estimates in the atomic bomb survivors. *Radiat Res* 114:437-66, 1988 (RERF TR 9-87)
16. Ban S, Awa AA, Ezaki H, Hiraoka T, Yamane M, Nishiki M, Dohi K, Miller RC, Bender MA, Setlow RB: Radiosensitivity and chromosomal aberrations of human skin fibroblast cell strains from atomic-bomb survivors with breast cancer. (RERF TR in preparation).
17. Hakoda M, Akiyama M, Hirai Y, Kyoizumi S, Awa AA: In vivo mutant T cell frequency in atomic bomb survivors carrying outlying values of chromosome aberration frequency. *Mutat Res* 202:203-8, 1988 (RERF TR 19-88)

原爆被爆者データの放射線被曝線量推定値における
確率的誤差に対する補正について

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

Donald A. Pierce, Daniel O. Stram, Michael Vaeth

放影研業績報告書シリーズ

REF Technical Report Series

業績報告書は、放射線影響研究所で行われたオリジナルな研究を報告するためのものである。このシリーズに含まれる報告書は、研究所内外の専門家による検討を受けたものであり、その一部又は全文は学術雑誌への投稿原稿としても用いられる。本報告書を引用したり、参考文献として言及しても差し支えないが、本報告書自体は当研究所の内部出版物である。報告書の入手を希望される場合は、〒732 広島市南区比治山公園5-2放影研編集課あてにお申し込み下さい。

1989年から、放射線影響研究所の業績報告書は、従来の日英両文を併記した方式では発行しない。主要な報告書については、今後も日英両文で印刷するが、それぞれ別に発行する。内容が高度に専門的であり、一般の関心が少ないと思われる報告書については英文のみとし、日本文の要約を添付する。

これにより、広島・長崎の原爆電離放射線被曝の人体に及ぼす晩発性生物学的影響に関する最近の知見を今までよりも速やかにお知らせできることと思う。

放射線影響研究所(元 ABCC)は、昭和50年4月1日に公益法人として発足した。その経費は日米両国政府の平等分担とし、日本は厚生省の補助金、米国はエネルギー省との契約に基づく米国学士院の補助金とをもって充てる。

原爆被爆者データの放射線被曝線量推定値における 確率的誤差に対する補正について[§]

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

Donald A. Pierce, Daniel O. Stram, Michael Vaeth

統計部

要 約

原爆被爆者の放射線被曝個人線量推定値に非系統的誤差が存在するため、線量反応解析において放射線の影響が過小評価され、線量反応推定曲線を歪める結果となっている。線量推定誤差に対して有効な統計モデルがあるという前提で、線形及び二次線量反応モデルの両方について、このような偏りを修正する統計学的方法をここに示す。このような誤差の性質と大きさについては、現在のところ期待されるほど分かっていないのでこの前提は重要である。ここでは、どちらかといえば微妙な統計学上の問題を明確にすることに重点を置いた。この方法は線量推計値の下方修正を含むが、このことはこれらの推定値に偏りがあることを意味するのではなく、むしろリスク推定における偏りを除くための線量反応解析の一部をなすものである。本報告書は主要な焦点を線形線量反応モデルにおいているが、線形・二次モデルの場合の方法も示す。被曝線量推定上の誤差について、信頼できそうな統計モデルを幾つか検討し、誤差の大きさの実質的な範囲を示し、得られた偏りの修正値について感度分析をした。これらの誤差モデルでは、癌死亡の線形過剰リスクの不偏推定値は、誤差を補正しない推定値の約5%~15%増であることが分かった。通常行われているように、極端に高い線量をもつ若干名の被爆者をこの解析から除くと、この幅の上限は約10%減少する。この範囲は極度に広くはなく、また難しさはあるが統計的問題はこの方法を進めるのに十分明確であるので、ほとんどの解析にここに提案する方法による補正を行うことを勧告する。そのような補正の暫定的なものとして、検討した方法のうち中心部分の範囲における特定の誤差モデルを一つ提案する。この勧告にはDS86新線量推定方式に特有の情報は用いられていないが、早急にDS86に関連する情報を組み入れることは重要であろう。被曝線量体系の改正は、主として被爆生存者の多数の集団に影響を及ぼす系統的な誤りを修正することを意図しているが、こ

[§]本報告の英語版は別に発行した。

に提案した方法は、個々の被爆者の被曝位置と遮蔽状況に関する不完全な情報、つまり、ほとんどが新旧両線量体系に共通な入力情報からもたらされる非系統的な誤差にのみ関与するものである。新線量推定体系の精度に関する今後の情報が極めて有用であるためには、これに関連した微妙な統計学的諸問題に注目することが極めて重要である。特に、真の被曝線量値に属する者についての推定値分布と、ある推定値を与えた者についての真の被曝線量値分布を区別することが肝要である。

緒言

放影研における調査研究のほとんどは、相当の不確定要素を含む放射線個人被曝線量推定値を利用して、原爆被爆者における各種指標の線量反応関係の推定を行うことである。今回の報告では、この被曝線量推定値につき、原爆の出力などに起因する系統的誤差ではなく、被爆者が原爆時にいた場所やその遮蔽状況にかかわる不確定要素を主原因とする「確率的誤差」を中心に検討を行う。線量反応解析の際にこの種の確率的誤差が「平均化」されないことは、統計学の一般的な原則からも、また、この誤差を扱った過去の研究結果からも明らかである。むしろ、そのために線形リスクや線量反応関係の形状の推定に系統的な誤差をもたらしている。本報告の目的は、この種の偏りを軽減する統計学的方法を示すことである。ここでは、単純化のために真の線量に対して線形関係を仮定したモデルを中心に考察を行い、勾配パラメーターの偏りを少なくする方法を考案した。しかし、この方法を拡張して、2次の項を含むモデルを用いた線量反応関係の推定にも応用することは容易であり、その場合の方法についても言及するが、詳細は Pierce 及び Vaeth¹ の報告に記述されている。

線量誤差の問題は、以前に Jablon² 並びに Gilbert³ によってかなり詳細に検討された。前者の報告は、被曝線量推定誤差の確率的な性質や大きさについて解析を行うとき特に有用である。今回の報告は、結論が後者の報告と極めて類似しているが、考察の細部や提案する統計学的方法に実質的な差がある。発癌リスクの増大を検討するときに興味の対象となる生存率データを解析する場合、線量誤差に関連して生じる特別な問題点については Prentice^{4,5} が考察を加えている。

本報告の主目的は、統計学文献の主流の中であまり目立つことのない微妙な一般的な問題、及び当該データ特有の問題の双方に対応するための接近方法とその論理的根拠を述べることにある。幾つかの具体的な結果も示すが、その結果も誤差の統計モデルに対していかなる仮定を設けたかによって左右されることは勿論である。ここでは、誤差の実際の性質について検討を試みない。しかし、若干の全般的な説明が必要であろう。

まず、誤差の性質について、上記の系統的誤差と非系統的誤差の関係を含めて、詳細に論じることは非常に複雑になる。ここで使用したような誤差の統計モデルは、問題の過剰単純化であることは確かであるが、もう一方の極端として、細部や複雑な内容を重視すれば、もっと多くの制限が加わることさえある。更に、最近に至って線量計算方式の大幅な修正が

行われた。⁶ DS86と呼ばれる新線量体系では、従来の線量計算方式に含まれていたかなりの程度の系統的誤差の修正が行われた。しかし、非系統的誤差の構造に影響を及ぼすような変更もある。DS86報告⁶の第9章に、新線量推定方式の精度の問題に関する予備的考察と、線量体系全体についての便利なまとめが記載されている。DS86推定値の精度が全体としてどの程度であるか、特に各遮蔽分類別、あるいはもっと細かい層化分類別にみた場合にどの程度であるかについて、一層の情報が必要である。より多くの情報が得られるにしたがって、ここに提案する接近方法も進化していくであろう。放影研及びその他の研究所においても、ここに示す統計学的方法の改良と、DS86推定値における誤差の実際的な性質と程度の解明に向けての研究に、高い優先性を与えるべきものとする。しかし、今の段階では、従来の線量計算方式に対して Jablon² が行った誤差の評価に引き続き大きく依存しておくことは妥当と思われる。同氏は、線量計算に用いられる各種の入力パラメーター（被爆者の位置、遮蔽の種類など）における不確定要素の程度の定量化を試み、この種の確率的誤差が線量推定値の不確実性にいかなる影響を及ぼすかを評価検討した。その主な論点は新線量体系にも当てはまるので、Jablonが提案したモデルを出発点として誤差モデルに用いる各種の仮定について感度分析を行った。

反応変数 y と説明変数 x との関係を求める場合の統計理論では、一般に y 変数の確率変動、すなわち、「誤差」に重点が置かれることが普通である。 x 変数も測定誤差を伴う場合については、もっと専門的になるが、膨大な文献がある。この種の文献から有用な背景情報が得られることは間違いないが、原爆被爆者データの特異性のために、想像されるほどには直接役立つ。古典的な文献の中で一般的に有用なものとして Cochran⁷ の報告があるが、最近開催されたワークショップ⁸ ではもっと新しい文献が紹介され、広範囲な教本が Fuller⁹ によって最近だされている。また、本報告に述べる技法と特にかかわりのある研究が Armstrong,¹⁰ Schafer¹¹ 及び Prentice^{4, 5, 12} によって行われており、Clayton¹³ がその再検討を行った。この最後の文献では、不正確な被曝量を使用した疫学的調査から得られるデータの統計的モデル化を行う場合の統合的な接近方法を記述しているが、その方法、特にこの一般的なモデルの解析に用いる計算方法は一層の改良が必要である。この領域での今後の発展によって本報告に示した方法にも改善すべき点が示唆されるであろう。

ここに示した統計的手法の要点は、(線形)線量反応解析に先立って被曝線量推定値に補正を加えることである。現行の、被曝推定値を用いてまず線量反応の解析を実施した後にパラメーター推定値を補正する場合に比べて、この方法が望ましいと考える理由は種々ある。この方法では、補正を加えることによって、低線量域を除けば、現行の推定値が非線形的に減少し、その減少の程度は推定被曝線量が高いほど大きくなる。このような操作を行うことについて若干の議論があることは、DS86体系の設定に傾注された努力を考えると当然である。そこで、この技法の展開に関する専門的な細部について記述する前に、かなり微妙な側面を有するこの基本的概念を説明する。

最も基本的な概念としては、次の三つである。

- i) コホートの中で真の被曝線量 x がほぼ同じである者の推定被曝線量 z の平均値。これを $\text{Avg}(z:x)$ として表す。
- ii) コホートの中で推定被曝線量がほぼ同じ者の真の被曝線量の平均値。これを $\text{Avg}(x:z)$ として表す。
- iii) 線形モデルにおける各パラメーター推定値の平均値、又は偏り。これは基本的な標本抽出変動と被曝線量推定値の確率的誤差の両面から平均化したものである。

上記の i) に関しては、いかなる線量体系でも、その合理的な目標は、原則として、すべての x に対して $\text{Avg}(z:x) = x$ という関係が保存されるということであろう。「原則として」と述べた主な理由は次のとおりである。すなわち、a) ある程度の合理的な近似としてこの関係が得られるだけで十分であること、並びに、b) 対数(又はその他の)目盛りを用いて平均化を行う場合、この関係が成り立てば満足であると考えられることである。ただし、その場合、本来の目盛りではその関係は成り立たない。上記の i) に関する一般的な議論は、仮に個人の位置や遮蔽状況などの入力パラメーターに起因する誤差を除けば、線量計算方式は正確な推定値を与えているという理想的な場合を考えてみればよいであろう。その観点では、入力パラメーターとして適当な推定値を用いれば、特に対数目盛りを使用するならば、少なくとも合理的な近似として $\text{Avg}(z:x) = x$ が導かれる。

しかし、前記の各文献で論じられているように、このことは線量反応関係の推定に用いる各パラメーターの推定値に不偏性を保証するものではないことは基本的な統計結果である。この点が iii) における問題である。推定被曝線量 z に基づいて線形リスク推定を行う場合の偏りは、上記 ii) の平均値 $\text{Avg}(x:z)$ と z との間どの程度の差があるかということに直接関係がある。この点に関しては、「偏り」が確かに存在し、推定被曝線量の高線量域では $\text{Avg}(x:z)$ が z よりもかなり低いことを本報告で示す。更に、この種の偏りを補正することは、ここで関心の中心となる偏り、すなわち、パラメーター推定値における偏りを除去する上での鍵になることが明白になるであろう。

$\text{Avg}(x:z)$ が z よりも系統的に低いということは、線量体系に「偏り」があることを非難する意味ではない。実は、この現象を避けた線量体系を設けようとする自体に無理がある。特に、少なくとも当調査研究下での主目的の一つとなる上記 i) の意味での偏りの除去は、ii) 及び iii) の意味での偏りの除去とほとんど相容れないものである。原爆被爆者データでは、高線量になるほど人数が急激に減少し、そのために推定被曝線量の高線量域では $\text{Avg}(x:z)$ が z よりもかなり低くなる。「偏り」という言葉が微妙なものをさす専門的な意味をもつと同時に、日常的には強い意味をもつ言葉として用いられることは残念である。

この問題に対処する統計的方法としては、線量反応解析の際に、推定被曝線量 z の代わりに補正推定値 $\text{Avg}(x:z)$ を代入することが必要である。このような接近方法は線形(相対的)リスク・モデルの場合に有効である。これは線量体系における「欠陥」の修正を意図したものでないことが上記の説明から理解できよう。この点については、被曝線量推定値に補正を加えることを線量計算の一部としてではなく、むしろ線量反応解析の一部と考えることが混乱を避ける最も簡単な方法であろう。この考え方の妥当性を示すものとして、線量反応解析の際に、被曝線量推定の誤差に対処するために考慮すべき修正がほかにもあることを挙げるができる。例えば、線量反応関係の2項モデルの解析を行う場合、反応の変動に追加項目がでてくるなどである。なお、 z の代わりに $\text{Avg}(x:z)$ を代入するというこの単純な解決策は、非線形モデルの場合には有効でない。

表1に示す架空的な事例によって、この基本的な概念の幾つかを定量的に示す。各線量区分群に用いた被曝者数は、放影研の寿命調査コホートにおける人数とほぼ同じであるが、この例では現実的な誤差モデルを意図していない。真の被曝線量及び推定被曝線量を Gray (Gy) 単位で四捨五入した上、推定値が真の被曝線量と一致する確率を0.5、いずれかの方向に1 Gyの誤差がある確率を0.25とする統計的誤差モデルを仮定する。表1では、真の被曝線量と推定被曝線量との相互対照分割表に被曝者の期待人数を示している。推定値に、前記の i) に示した通常の意味での偏りがあるか否かの考察は、この表の行に關係するものである。他方、 $\text{Avg}(x:z)$ は列の平均値に關係するものであり、表の下段に幾つかの数を選んで示した。前記の ii) における意味での「偏り」は、推定値が真の被曝線量に比べて1 Gy 高い者よりも1 Gy 低い者の方がはるかに多いことがその原因であることが分かる。

表1に示した誤差分布を例として選んだ理由は、主として単純化のためである。しかし、これから幾つかの重要な点が明らかになる。誤差に起因する問題の主なものとしては、一般に次のように考えられている。すなわち、i) 被曝線量が高いほど誤差が大きいこと、また、ii) 対数目盛りを使用するとかなりの左右対称性を示すので、過小評価よりは過大評価が

表1. 基本的な概念を示す架空事例

	推定被曝線量						被曝者数
	1	2	3	4	5	6	
真の被曝線量	250	500	250				1000
3		75	150	75			300
4			33	66	33		132
5				15	30	15	60
$\text{Avg}(x:z)$			2.50	3.62			

著しくなることが考えられる。この二つの要因が関係してくるとはいえ、当事例にはそのいずれも含まれていないので、たとえその二つが存在しない場合でも、根本的な問題が残ることが表1から分かる。

本報告で使用したもっと現実的なモデルに基づく $Avg(x:z)$ 推定の方法の原理は、表1における計算方法に非常に近い。コホートにおける推定被曝線量の分布と誤差モデルから、被曝者の真の被曝線量の分布を推定することができる。これは原則として表1における右端の列に相当する。次に、統計的誤差モデルから z に対応する x 別の被曝者分布を求めることができるが、これは表の中央部に相当する。次に、このようにして求めた分割表から、前記の場合と同様にして $Avg(x:z)$ が推定できる。

材料及び方法

本章は、三つの部分から成り立っている。第1部では、誤差の統計的モデル、並びに補正被曝推定値として用いる $Avg(x:z)$ の推定方法について説明する。第2部では、補正した推定値を線量反応解析で使用する理由とその方法、並びにそれに関係する統計手法の詳細を説明する。最後の部分では、例として示した幾つかの線量反応解析に使用したデータを説明し、いかなる誤差モデルを選んで使用するかによって最終結果がいかに敏感に変化するかを示す。

$Avg(x:z)$ の推定

必要とする数量 $Avg(x:z)$ は、コホートの中で z の値が(ほぼ)同じである者の x の条件付き分布、これを $f(x:z)$ として表す、の平均値である。 $f(x:z)$ から計算されるその他の数量、例えば $Avg(x^2:z)$ や標準偏差 $SD(x:z)$ などにも関心がもたれるであろう。この分布 $f(x:z)$ の概念は、ここでの展開の中心となるもので、表1を注意深く検討すればその意味がよく分かるであろう。ここで使用する確率要素は、個人の真の被曝線量を確率変数と考えたり、又はコホートを集団からの抽出標本と考えたりすることに基づくものではない。それは、被曝推定線量の誤差が確率的であると考えることによって導かれたものであることを指摘したい。これはデータを階級別に区分してみた場合に最も具体的に分かる。誤差を確率的に割り当てていけば、分類 z に属しているサブ集団の真の被曝線量は、 x の各区分すべてにわたって確率的に分布することになる。

このような階級別区分によって具体性が与えられるが、被曝線量を区分しない方法の開発を進めることにも重要な利点がある。中でも最も重要な点は、表現が簡単で明瞭になること、並びに特定の解析に最適と思われる階級区分にその結果を適用できることである。しかし、厳密な被曝線量を使用しての方法論の開発を進めることの意味は、後にその方法を階級別に区分したデータに適用して結果を導くことができるからと考えることの方が一番よいかもしれない。

特定の被曝推定線量が与えられた者における真の被曝線量の分布の密度関数 $f(x:z)$ は次式から求められる。

$$f(x:z) \propto f(x) f(z:x) \quad (1)$$

ここで $f(x)$ はコホートにおける真の被曝線量の分布、 $f(z:x)$ は推定誤差について仮定したモデルの密度関数である。ここでの比例関係は x に対してであるので、 $f(x:z)$ の積分が1になるように各 z の比例定数を決定すればよい。上記の $f(x:z)$ の定義によれば、 $f(x)$ はコホートにおける真の被曝の分布であって、コホート(被曝線量)がそこから抽出された仮定の被曝分布ではないことを強調したい。 $f(x)$ と $f(z:x)$ の選び方について下記に説明する。本報告書では、必要とする数量 $\text{Avg}(x:z)$ の計算は数値積分によって行うので、モデルの選択は数学的な便宜のための人為的な制約を受けない。ここで使用した誤差モデルの場合、 z の値を適当に選んで $\text{Avg}(x:z)$ を計算してみると、簡単な経験的数式から「低減因子」として $[z - \text{Avg}(x:z)] / z$ の適合度が非常によいと認められた。いずれの誤差モデルを用いる場合も、線量反応解析の際にこの数式を使用して被曝推定線量が補正できる。実際は、データの分割表における各セルの平均被曝線量にこれが適用されることが多いであろう。

したがって、推定誤差についてあるモデルを仮定したとき、コホートの真の被曝線量の分布 $f(x)$ が分かれば、 $\text{Avg}(x:z)$ が直接計算できる。両市の間で $f(x)$ が異なるので、² $\text{Avg}(x:z)$ の計算は都市別に行う。この都市別の分布 $f(x)$ として単純に推定線量 z の分布を採用しても近似値としてそれほど悪いことはない。しかし、ある誤差モデルを仮定して何らかの補正を加えて $f(x)$ のもっとよい推定を求めることは容易である。この点について今回の研究で採用した接近方法は、結果の章でその概要を説明する。

両市における推定線量の分布を表2に示した。ここで使用した単位は厳密な意味で「線量」でなく、DS86カーマ(kerma)である。すなわち、これは最近導入されたDS86線量体系⁶に基づくものであり、被曝者がいた場所における構造物や地形による遮蔽の影響を補正した空中組織カーマである。本報告書における方法の開発はすべてカーマに基づいて行った。放影研

表2. 被曝線量が特定のカーマ値(Gy)を超える被曝者数

被曝線量	広島	長崎	被曝線量	広島	長崎
0.1	14632	3420	3.0	338	129
0.5	5099	1501	3.5	218	100
1.0	2229	801	4.0	171	85
1.5	1311	444	4.5	128	71
2.0	820	282	5.0	108	55
2.5	532	185			

における解析では、人体による遮蔽を考慮した臓器線量を用いることがある。その場合、上記の低減因子をカーマとして計算し、次いで推定臓器線量にそれを当てはめることを勧告する。それは計算を簡単にすることが主な理由である。臓器線量をもっと直接的に補正するためには、各臓器ごとにここで述べるような計算が必要である。臓器線量は、カーマに比べて確率的誤差が若干大きいことを指摘したい。これについては、DS86報告⁶の第9章を参照。

コホートには、比較対照群として被曝線量が実質的にゼロである者が多数含まれている。ここで直接関心の対象となるのは真の被曝線量がゼロ以外の者だけであるので、比較対照群の中で推定線量がゼロと計算されている者は、その真の被曝線量がゼロに近似すると仮定して無視することができる。なお、被曝線量推定値を6 Gyで打ち切ることが慣習となっているが、それは真の被曝線量がこれ以上になれば恐らく生存の可能性がほとんどなく、したがって、そのような高線量推定値は信頼性が特に悪いと考えられるためである。線量反応関係の解析を行う際、この種の何らかの処理方法を継続することが望ましい。しかし、今回の基本的な方法の開発ではこのような打ち切りは無視したが、これは、特に被曝線量推定が極めて高線量の場合、誤差モデルの妥当性についての懸念を軽減する効果がある。

次に、ある特定の真の線量に対応する推定線量が示す分布について、いかなる統計モデル $f(z|x)$ を選ぶべきかに注意を向ける。ここで中心となるのは Jablon² が提案したようなモデルである。それは誤差の主要出所源についての考察が主な契機となって提案されたもので、 $\log(z)$ が平均 $\log(x)$ 及び x とは独立的な標準偏差をもつ正規分布であると仮定していた。この種のモデルでの要点は、本来の日盛り上で誤差が真の被曝線量のレベルに比例することである。これは非常に妥当であると思われ、いかなる代替モデルにおいても恐らく採用すべきものであろう。ここで関心の対象となる値の範囲内では、 $\log(z)$ の標準偏差 (SD) は z の標準偏差の平均値との割合、すなわち、変動係数 $CV(z)$ にほぼ等しい。例えば、 $SD[\log(z)]$ の値 0.30, 0.35, 0.40 は、 $CV(z)$ の値 30.7%, 36.1%, 41.7% にそれぞれ相当する。単純化して、例えば、 $SD(z) = 0.30$ である場合のこのモデルを「対数正規30%誤差」モデルと呼ぶことにする。このような対数正規モデルでは、 $CV(z)$ が $\exp[SD(z)^2] - 1$ の平方根に等しいことを指摘したい。普通は $CV(z)$ として近似式 $\exp[SD(z)] - 1$ が用いられることが多い。しかし、これは $SD(z)$ がここで使用した値よりも小さい場合においてのみ正確である。

この種のモデルのもう一つの特徴は、誤差が対数日盛り上で左右対称的に分布していることであり、したがって、本来の日盛り上では過剰評価の方向に偏っている。このことは、真の被曝線量の対数が原爆からの距離とほぼ線形関係にある事実から分かる。上記のように、このモデルに基づく推定値に偏りが無いのは対数日盛りの場合のみである。しかし、純粹にこの種の対数日盛りとの関係で本来の日盛り上で起こる偏りは、ここで検討する第一義的な問題ではあまり重要でない。

Jablon²は、対数正規30%誤差モデルが特に考慮に値すると結論したが、これよりもやや大きい誤差を仮定したモデルも真剣に考慮すべきであるとした。ここではそのモデルを出発点としたが、もっと大きな誤差やその他のモデルを用いたときの影響の感度分析に重点をおくため、対数正規40%誤差モデル、並びに対数正規ではないがCV(z)がほぼ40%であるモデル2種類の結果をも示した。対数正規35%誤差モデルについても幾つかの最終的な数値結果を示した。

対数正規モデルに関する最も重大な懸念は、対数目盛り上で左右対称性を仮定したことにあるのではなく、正規分布というもっと具体的な仮定を設けたことにある。正規分布は非常に「裾が軽い」ことが特記されるべきもので、実はそのために悪評がある。すなわち、それは典型的な事象において起こり得るかなり大きな偏差には対応できないことである。今回の結果が、極端な誤差が起こり得る確率に対して敏感であると想像できるので、「もっと重い裾」をもつ分布の検討を行うことも重要である。当然な選択として、汚染正規分布と呼ばれるものを用いて、 $\log(z)$ が標準偏差0.30をもって正規分布する確率を85%、標準偏差0.75をもって正規分布する確率を15%としてモデル化した。以上のパラメーターは、上記の第2のモデルのように $\log(z)$ の標準偏差が0.40になるようにして選んだものであるが、その分布の形状は全く異なる。

対数目盛り上で左右対称性を示すであろうという仮定の不成立とその結果生じる過大評価の方向へのゆがみを検討するために、第4のモデルを選んだ。このモデルではzは正規分布し、平均値x、及び $CV(z) = 40\%$ とすることによって先の二つのモデルとの比較ができる。対数目盛り上でほぼ左右対称であると考えることがもっと自然であると著者は思うが、この第4のモデルはこれとかなり極端に離れたものになる。しかし分布の形状による影響については有用な情報が得られる。

したがって、次の4種類の誤差モデルに基づく結果を示す：

- A. 対数正規 30%誤差
- B. 対数正規 40%誤差
- C. 汚染対数 40%誤差
- D. 正規 40%誤差

無遮蔽、日本家屋内、重構造ビル内、など各遮蔽分類別に、それぞれ特異的な誤差モデルを使用すべきか否かを検討することは自然である。今回ここに開発した技法の延長として、これらの因子に応じて、あるいは恐らくその他の因子にも応じて、%誤差を変更できるようにすることが、最も重要となるであろう。DS86線量体系における今後の発展は、今回の方法にこの種の改良を行うべきであるかどうかを評価するときに必要な情報を与えるであろう。

線量反応解析における補正値の利用

まず、期待される反応が真の被曝線量と線形関係にある場合のモデルを考える。ここで対象となるデータの種類や統計技法に対して、なぜ推定被曝線量 z の代わりに補正推定被曝線量 $\text{Avg}(x:z)$ を用いればパラメーター推定の偏りが減少するか、その理由について下記に示す。非線形モデルについては、本章の終わりに述べる。ここでの「偏り」の定義は、通常の標本抽出変動と推定被曝線量における確率的誤差の双方について平均化を行うことに関係する。上記の置き換えによって偏りが減少するとはいえ、数量 $\text{Avg}(x:z)$ が真の x と同値であると考えるべきでないことは明らかである。したがって、解析あるいは解釈の際に、これ以外に補正が必要であるか否かを検討することも重要である。ここでは要点についてのみ焦点を当てるように努め、技術面の細部は付録に示す。

考察を更に進める前に、コホート全体の解析を行う場合も、推定被曝線量に基づいて抽出したサブ集団の解析を行う場合も、同じ補正値 $\text{Avg}(x:z)$ を使用すべきであるということを目指したい。被曝コホート全体における真の被曝線量の分布 $f(x)$ が計算に関与しているためにこのことが問題となり、サブ集団を用いる場合は、この分布をその特定のサブ集団に限定して適用すべきであるとの考えもできよう。しかし、標本抽出は x 値に基づいて行うものではなく、必ず z 値に基づいて行われるので、その考えは妥当ではない。分布 $f(x:z)$ は z の選定方法に左右されないので、ここで開発した方法は、選択された集団の解析を行う場合にも、そのまま適用できる。

下記に示す技法の開発に当たっては、共変数として被曝と都市のみをとる。これ以外の測定誤差を伴わない共変数を追加導入しても、基本的な結果は変わることはない。ただし、その共変数は、真の被曝線量の分布あるいは推定誤差と関連があってはならない。寿命調査の場合、関心の対象となる主な共変数は都市、性別及び被曝時年齢である。都市の影響は、前述のように補正値 $\text{Avg}(x:z)$ を都市別に計算することによって処理する。原則として性、被曝時年齢別の補正も考えるべきであろう。しかし、モデルの二つの構成部分のいずれにも性別又は被曝時年齢によって本質的な差がないので、その補正を加えても効果は軽微であろう。

標準的な解析方法に対して、補正推定被曝線量を代入する以外にその他の変更を加える必要があるか否かは、解析されるデータの種類の依存することが確められた。放影研で関心の対象となっている次の2種類の主要データについて考えてみよう。すなわち、i) 癌発生率、又は死亡率。この場合は生存率解析の統計方法を使用する。ii) 染色体異常有病率。これには2項分布データの解析方法を用いている。前者の場合、標準的な統計方法や解釈に対してこれ以上の変更は不必要である。しかし、後者の場合は、補正線量を用いて2項分布モデルに基づく解析を行うだけでは不十分である。

この二つの間の違いは、通常の標本抽出変動と推定被曝線量の誤差に起因する変動との相対的な割合に関係している。i) の場合、各対象者のデータは本質的に2値性であるのに

対し、ii)の場合のデータは多数の被検細胞中における異常細胞の割合である。後者では、通常の標本抽出変動は各対象者について検査した細胞数に反比例するので、前者に比べてはるかに小さい。

癌データでは、癌発生、又は癌死亡までの期間が確率変数であるとともに各対象者のデータの一部でもあるので、もっと複雑な側面が若干ある。しかし、Coxの回帰分析法を用いて線形相対モデルに基づいて癌データの解析を行う場合、 z の代わりに $\text{Avg}(x:z)$ を代入するだけで十分であることが Prentice⁴ によって示されている。癌データの解析に放影研で使用している統計モデルは、本質的にはこの方法の時間階級区分法であり、データを区分した形で具体的に示すことによって Prentice の結果がもっと明瞭になるであろう。その詳細は付録に示した。

染色体データの場合、2項分布モデルは極めて大きな分散を伴うことは明白であり、かなり以前から2項分布法に修正を加えてデータの解析が進められている。例えば、Prestonら¹⁴の報告を参照。この種の特別な方法(付録で説明している)を利用する限り、補正推定被曝線量の代入を追加するだけで十分である。

真の被曝線量からみれば、被曝線量 x を有する対象者の観察結果は次のような形のモデルに従う。

$$y = \alpha + \beta x + \text{誤差} \quad (2)$$

ここでは、 y は対象者から求められた細胞標本中の異常細胞の割合である。このデータに適用する標準的な最大尤度法は、このモデルに基づく加重最小自乗解析と考えることができる。「誤差」の期待値はゼロであり、これは不偏な推定を意味している。「誤差」の分散は x に依存するので、加重回帰分析が必要となる。また、 α 及び β にも依存するので、反復法が必要である。

コホートの中から無作為抽出した推定被曝線量 z を有する被曝者に対して、(2)は次のように書き改めることができる。

$$y = \alpha + \beta \text{Avg}(x:z) + \beta [x - \text{Avg}(x:z)] + \text{誤差} \quad (3)$$

この場合、 x は分布 $f(x:z)$ に従う確率変数であるので、 $\beta [x - \text{Avg}(x:z)]$ の項は、平均ゼロ、分散 $\beta^2 \text{Var}(x:z)$ の「誤差項」としてモデルに追加されることに注意せねばならない。したがって、ここでの最後の因子、すなわち、特定の z を有する者における x の分散は重要な数量となり、 $\text{Avg}(x:z)$ と同様、特定の誤差モデルから計算できる。

したがって、初等的結果として推定被曝線量 z の者におけるデータは次のように示すことができる。

$$y = \alpha + \beta \text{Avg}(x:z) + \text{誤差}^* \quad (4)$$

ここで誤差*は(3)における二つの誤差項の組み合わせである。標本抽出変動と被曝線量推定の誤差の双方を平均化したときに誤差*の期待値がゼロであるということは、(4)に基づく加重線形回帰分析によって不偏な推定値が得られることを示している。なお、標準的な計算から次が導かれる。

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

ここでの最後の項は、(2)の誤差の分散の分布 $f(x:z)$ による平均である。したがって、(5)は加重最小自乗法を用いて(4)のフィッティングを行うための適当な加重値を与えているとともに、最終的にはパラメーター推定の標準誤差を与える。

癌データの場合、(5)の最初の項は消失する。すなわち、(2)における誤差の分散が x の関数であったと同様、(4)における誤差*の分散は $\text{Avg}(x:z)$ の関数である。したがって、 z の代わりに $\text{Avg}(x:z)$ を代入する以外は、解析方法に変更を加える必要はない。染色体データの場合、(5)におけるこの二つの項の大きさはほぼ同じである。これらの結果の詳細については付録に記述したが、そこでも示したように、染色体データに適用されてきた特別な方法は(5)の形の分散にも適当であると認められ、補正線量推定を代入する以外に変更の必要はない。最後に、(2)の場合と異なり、加重最小自乗法はもはや最大尤度法でないにもかかわらず、その方法を用いる理由を付録に示した。

上記のすべての結果は、期待される反応が真の線量に対して線形であると仮定したモデルに基づいている。これに対応する非線形モデルの開発は一般にもっと困難であるが、線形・2次曲線モデルへの延長、すなわち、(2)における $\alpha + \beta x$ を $\alpha + \beta x + \gamma x^2$ に拡張することを考えても不自然ではない。この場合上記の延長として z^2 による解析の代わりに $\text{Avg}(x^2:z)$ を代入することにより、偏りのない推定が得られる。なお、数量 $\text{Avg}(x^2:z)$ は、 $\text{Avg}(x:z)$ の場合と同様の方法で計算できる。その計算が必要なことは、実は $\text{Var}(x:z)$ 、すなわち、 $\text{Avg}(x^2:z) - [\text{Avg}(x:z)]^2$ の形で既に生じていた。この接近法の採用は、特に慎重を要する。すなわち、 $\text{Avg}(x^2:z)$ は、いかなる誤差モデルを選ぶかによって、 $\text{Avg}(x:z)$ よりも敏感であると予想される。被曝推定値が誤差を伴う場合の線形・2次曲線モデルに基づくパラメーター推定については、癌死亡率の線量反応関係の形状に関する Pierce 及び Vaeth¹ の最近の報告に詳細に記述されている。なお、非線形モデルを用いる場合、前記の被曝線量推定値の誤差に起因する変動について若干の補足が必要となる。これらの点については、以下の節で説明する。

感度分析に用いたデータ及びモデル

ここで採用した誤差モデルの感度を分析する最も有効な方法は、その誤差モデルに基づく補正が、実際の線量反応解析におけるパラメーター推定にいかなる影響を及ぼすかをみることである。そこで、Preston 及び Pierce¹⁵ が最近の解析で用いたデータとモデルをそのまま使用して癌死亡率データを解析してみた。白血病を除く全癌を一群として、また、白血病だけを一群として、線量と線形関係を仮定したモデルを用いて過剰相対リスク推定値を計算した。このモデルでは、年齢別過剰相対リスクは年齢に対して一定で、性別及び三つの原爆時年齢群に依存すると仮定した。新旧線量の比較に当たって Preston 及び Pierce¹⁵ が行ったと同様、対象となった6群の単位被曝線量当たりのリスクの単純平均を総括的リスク推定値とした。白血病以外の癌では、身体によって臓器がかなり遮蔽されていた線量の代表値として、大腸の被曝線量を用いて解析を行った。白血病では骨髄線量に基づいて解析した。なお、線量計算方式によって与えられる臓器線量推定値については、カーマとして計算した低減因子を適用して補正を行った。

結 果

まず、両市における $f(x)$ の推定に考察を加える。被曝線量推定が正の値である被爆者の中で、いずれかの z の値を超える者の割合に対しては、Weibull モデルの分布関数である $\exp(-\theta_1 z^{\theta_2})$ の形の適合が良好である。(θ_2 の値は1よりも小で、単調に減少する密度関数に相当する。) そこで、 x の分布についても同様の形のパラメーターを仮定し、次いで表2に示したデータを用いて、かつ、被曝線量の推定誤差の影響を補正して、そのモデルのパラメーター θ_1 及び θ_2 を推定するという方法が適切であると考えられる。この実行は、かなり簡単である。すなわち、ある誤差モデルが与えられたとき、(数値計算によって)誘導された z の理論的分布と観察された分布とが極めてよく一致するように各パラメーターを選択(直接検索)する。これらの解は、上記で定義したいずれの誤差モデルにもあまり大きく依存しないので、以後は単純化のためにすべて第5の「中間的な」モデル、すなわち、対数正規35%誤差モデルに基づく値を用いる。その結果は次のとおりである。

広島	$\theta_1 = 2.84$	$\theta_2 = 0.50$
長崎	$\theta_1 = 2.33$	$\theta_2 = 0.50$

θ_2 が両市で等価であるのは、そのように決めたためではない。これは偶然であったが、当てはめた値が小数点第2位まで互いにほとんど一致し、ここでわずかに四捨五入したにすぎない。このようにして選んだパラメーターによって得られる適合度の精度は、ここでは詳細に述べないが、非常によいと認められた。

$f(x)$ についてこのような都市別モデルを使用すれば、 $\text{Avg}(x:z)$ の数値計算は容易である。選択したそれぞれの z 値につき、各誤差モデル別に計算した結果を表3に示した。

表3. 4種類の誤差モデルに基づく補正被曝線量推定値

被曝線量推定値 (Gy Kerma)	Avg(x:z)							
	広島				長崎			
	A	B	C	D	A	B	C	D
0.5	0.50	0.50	0.50	0.51	0.51	0.51	0.51	0.53
1.0	0.96	0.94	0.94	0.97	0.98	0.97	0.97	1.00
2.0	1.84	1.73	1.75	1.81	1.89	1.82	1.83	1.88
3.0	2.66	2.45	2.47	2.59	2.75	2.59	2.62	2.71
4.0	3.44	3.12	3.13	3.34	3.58	3.32	3.36	3.50
5.0	4.20	3.75	3.74	4.07	4.38	4.01	4.05	4.26
6.0	4.93	4.35	4.28	4.78	5.16	4.67	4.69	5.01

A: 対数正規 30%誤差; B: 対数正規 40%誤差; C: 汚染対数正規 40%誤差; D: 正規 40%誤差

前節では、 $\text{Var}(x:z)$ と $\text{Avg}(x^2:z)$ 、すなわち、前者は特定の z の値を有する者における x の分散であり、後者は $\text{Var}(x:z) + [\text{Avg}(x:z)]^2$ で与えられるものであるが、この二つに関心をもたれる理由の幾つかを説明した。この数値結果は変動係数の自乗として、すなわち、 $\text{Var}(x:z) / [\text{Avg}(x:z)]^2$ として示すことが最も簡単であり、4種類の誤差モデルに基づくそれぞれの値を表4に示した。

表4. 4種類の誤差モデルにおける変動係数の自乗

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

A: 対数正規 30%誤差; B: 対数正規 40%誤差; C: 汚染対数正規 40%誤差; D: 正規 40%誤差

次節で表3、表4の結果について考察を加えるが、ここではモデルCとモデルBがほぼ同じ結果を示したと述べるにとどめる。誤差モデルCとモデルDについてはこれ以上の検討は加えない。以下に示す数値結果は、対数正規モデルA及びモデルB、更にその中間的な対数正規35%誤差モデルのみに基づくものである。

この応用に当たっては、 $\text{Avg}(x:z)$ に関する便利な公式が必要である。減弱因子 $[z - \text{Avg}(x:z)]/z$ に対しては、 $\log(z)$ の2次の多項式の適合が非常によいと認められた。zから $\text{Avg}(x:z)$ を計算するとき、これらの近似値を使えば被曝線量0.5～6 Gyの範囲では0.01 Gy以内の精度がある。被曝線量推定値が0.5 Gy未満の場合は補正を加えないこと、また、6 Gy以上の場合には6 Gyに切り下げ、それを被曝線量推定値として補正の対象とすること（これは実際の推定値に補正を加え、次いで6 Gyに相当する補正值まで打ち切ることと同じ）を提案する。減弱因子の近似値を求めるための係数を表5に示した。

表5. 減弱因子計算のための係数

	対数正規誤差モデル パーセント誤差	各項の係数		
		定数	$\log(z)$	$[\log(z)]^2$
広島	30%	0.03597	0.05807	0.01166
	35%	0.04732	0.07623	0.01336
	40%	0.06036	0.09684	0.01314
長崎	30%	0.01500	0.05304	0.00885
	35%	0.01900	0.06545	0.01374
	40%	0.02817	0.08031	0.01558

表4に示した値に対してもこのような多項式の適合は良好であるが、それらの値は各zに対して非常に一定しているので、ある与えられた誤差モデルを使用した場合に変動が少ないことを強調すれば、ここで使用したモデルの過大解釈となる。表にみられるように、値が小さく、0.09～0.15の範囲にあることは重要である。染色体異常のデータの場合、被曝線量の推定誤差のために追加されてくる変動がこれらの数値に依存していることは、上記及び付録での詳細な説明から明らかである。なお、表4に示した値に1.0を加えると $\text{Avg}(x^2:z) / [\text{Avg}(x:z)]^2$ の値が得られ、これは真の線量との関係に2次曲線モデルを当てはめる場合に関連するものである。最後に、表4における値の平方根は、ある与えられたz値に対するxの分布の標準誤差を $\text{Avg}(x:z)$ の割合として示したものである。これについては、被曝線量推定値の解釈を行う場合、一般にある程度の関心がもたれる。

分布の標準偏差は、分布が正規分布でなければ、又は分布の形状に関するその他の情報が無い限り、変動の全体像を示すものではない。この点については、分布 $f(x:z)$ の幾つかの例をグラフとして図1に示した。同図では、対応する分布 $f(z:x)$ のグラフも示して、この二つの重要な分布の違いを明確にした。横軸は、一方の図でxに相当し、他方の図でzに相当することを指摘したい。

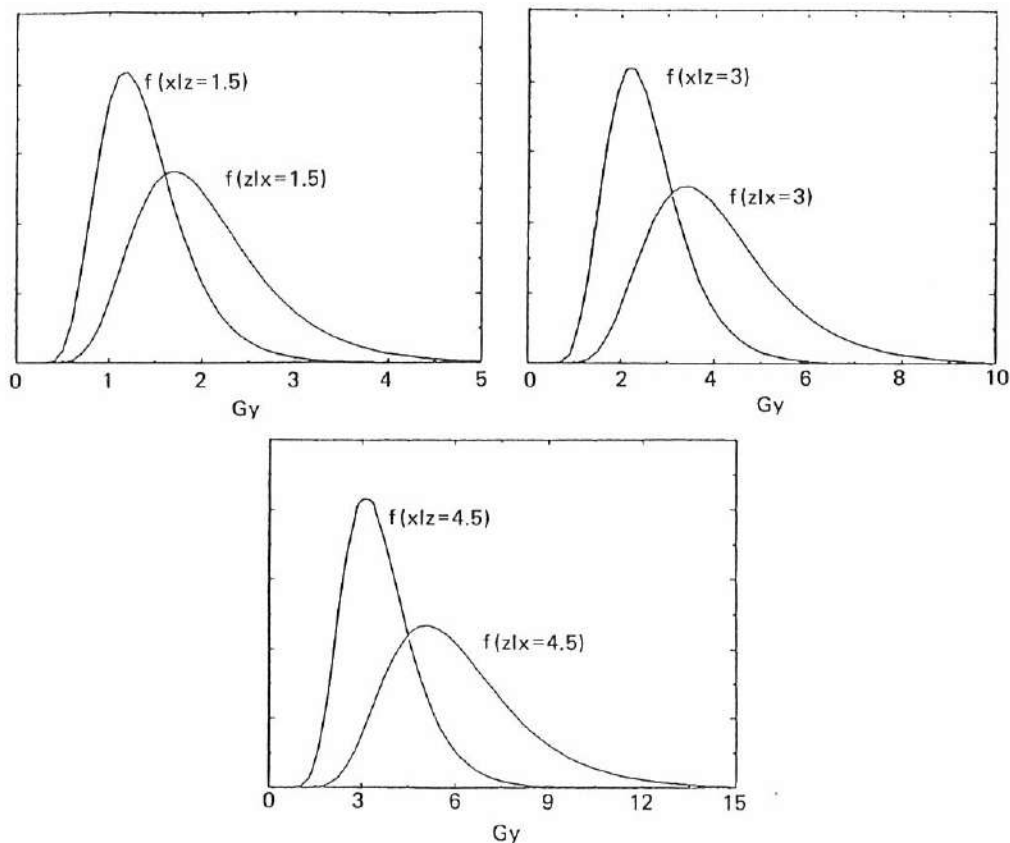


図1. 対数正規35%誤差モデルに基づく広島での $f(z|x)$ 及び $f(x|z)$ のグラフ

上記の $\text{Avg}(x:z)$ の値を各誤差モデル別に比較検討しても、実際的な意味で感度分析の明確な様相を知ることはできない。それは、線量反応モデルを解析する場合に線量の各部分の相対的重要性が不明（かつ説明困難）であるためである。そこで、表6に実際の線形線量反応解析に各誤差モデルを用いたときの影響を示した。データ、モデル及びパラメーター推定の要約方法については前節で説明した。要約したパラメーター推定値にみられる相対的な増加を示したが、これらのパラメーター推定値は、最近行われた線量体系の改定の影響を検討するため、Preston 及び Pierce¹⁵ が取り上げたのと同じものである。彼らの報告と同様、2種類の推定値を示したが、両者の比較は今回の目的に深くかかわるものであり、後述する。その一方はカーマ推定値を6 Gyに切り下げて調査対象集団全体の解析を行った場合のものであり、他方は推定値が4 Gy未満である者に限定して解析を行った場合のものである。具体的にいえば、表6に示した百分率の増大は、Preston 及び Pierce¹⁵ の報告の表11と12における最後の二つの項目に対応する。

表 6. 3 種類の対数正規誤差モデルに基づく癌リスク推定の増加

	対数正規誤差モデル パーセント誤差	過剰相対リスク推定値の増加	
		0-6 Gy	0-4 Gy
白血病以外の全ての癌	30%	10.0%	6.8%
	35%	13.3%	9.0%
	40%	16.7%	11.4%
白血病	30%	6.1%	4.3%
	35%	8.1%	5.6%
	40%	10.2%	7.2%

癌データの場合は、線量推定の誤差を補正して解析を行っても、パラメータ推定値に対するその推定値の標準誤差、すなわち、変動係数に本質的な変化はない。このことは、この新しい手法では反応変数の分散に関するモデルに変更を加える必要がなく、推定線量値の補正のみを行えばよいことを反映している。他方、染色体データの場合は、補正線量を用いて解析を行うと、反応変数の分散に余分なちらばりの項を追加することになるので、2項モデルの場合に比べて標準誤差が大きくなると予想されよう。しかし、2項モデルにおける過剰なちらばりを補正する方法が既に利用されており、ここに提案する方法の場合の標準誤差も、現行の方法に基づくものとよく似ている。

考 察

初めに指摘したように、ここで開発した方法は Gilbert³ の方法と基本的な違いはないが、特定の線量区分を選んで広範囲にわたる計算を行う必要がないので、多くの目的への応用に適している。その原理は、Prentice^{4,5} が報告したものと一致するが、同氏は生存率データへの応用を特に取り上げたのに対し、ここでは寿命調査の特異的な状況の細部に重点をおいた。Clayton¹³ が示した全般的な接近法は、モデルの三つの構成部分、すなわち、線量反応モデル、被曝線量の誤差、及び被曝線量の分布に関して、推計を行う場合の問題に統合的な解決を求めようとしたものであり、理論的にある程度の長所もあるが、計算方法はまだ開発中の段階にある。今回の接近法では、この三つの側面を個別に考慮するので、計算上の問題はかなり軽減される。しかし、Clayton の全般的なモデルの進歩に伴って本法の改善も考慮すべきである。回帰分析に当たって説明変数が測定誤差を伴う場合の問題を取り上げた文献の大多数は、本来の被曝推定値を用いてまず解析を行い、次に得られたパラメータ推定値に対して補正を加える方法が中心になっている。典型的な方法では、真の被曝の分布について

特別な仮定を設ける必要があるが、これは寿命調査データの場合に不適當である。しかも、対象となる寿命調査集団中の各サブ集団にみられる線量の分布に合わせて、特定の研究ごとに異なるパラメーター補正の方法を必要とする。ここで採用した接近法では、少なくとも真の被曝線量に対して線形モデルが適當であれば、線量反応解析におけるパラメーター推定値の、被曝推定値の確率的誤差に起因する偏りを軽減する妥当な方法を提供していることは明らかである。つまり、この方法では、被曝線量の誤差に関する真の統計モデルが既知であれば、

この種の偏りが本質的に除去され、また、近似的な統計モデルの利用によって偏りの実質的な減少を果たすことができる。被曝推定値に補正を加える必要があることは、線量体系に欠陥があるという意味ではないこと、また、ここに示した方法は被曝推定値における非系統的誤差に対応するためのものであるに過ぎないことを改めて強調したい。

唯一の大きな問題点は、それはもちろん回避不可能であるが、被曝推定値誤差の統計モデルに幾つかの仮定を設ける必要があるということである。以下の考察では、この種のモデルを使用する場合の感度分析の解釈並びに非線形モデルを考慮する場合の幾つかの問題を主に取り上げる。

まず、線形線量反応の解析を行うという目的からみれば、表3及び4は、被曝線量推定誤差に基づく偏りの度合について誤った印象を与えるとともに、誤差モデルの選択に敏感すぎる事が強調されているかもしれない。表2でみられるように、高線量域では被曝者が比較的少なく、したがって、この二つの問題について結論を導くために表6がもっと有用な基盤を提供する。

表3及び4は、かなり「重い裾」を有する汚染対数正規モデルと対数正規モデルとの間に違いがほとんどないことを示している。したがって、汚染対数正規モデルは、もっと現実的なモデルとしてかなり魅力的であるかもしれないが、このような一層の複雑性を求める必要はほとんどないように思われる。変動係数を一定とする正規モデルは、補正が少ないという意味で、幾分異なる結果を与える。このモデルについて、図1の場合と同様にして $f(x:z)$ のグラフを作成すれば、分布の右側の裾は図1と極めて似ているが、左側が著しく短い。これは、このモデルの予想外な一面を反映するものであり、それを理解することによって、対数正規誤差モデルの(線量の対数を尺度としたものではなく、線量を尺度とした)右側の裾が長いことについての懸念が軽減されるであろう。変動係数が一定である正規モデルでは、(1)における $f(x:z)$ 、つまり z を固定したときの x の関数の右側へのゆがみは、対数正規モデルの場合に比べてもっと大きい。すなわち、統計的にいえば、正規モデルの場合、 x が与えられたときの z の分布は左右対称的であるが、 z が与えられたときの x の「尤度関数」は右側へ非常に偏っている。したがって、この二つのモデルの間に実は微妙な差がある。筆者及びその他の研究者は、この問題を検討した結果、一般的な統計学的な理由と線量体系における特有な理由から、対数正規モデルの方がやや望ましいと考える。

表4から幾つかの極めて重要な情報が得られる。表4の数値が z にほとんど依存していないことは、少なくとも二つの理由で重要である。まず、染色体データにおける過剰な変動に対して以前から利用されている特別な補正方法は、ここでも適当であることを示している。第2に、 z に依存しないことは、 $\text{Avg}(x^2; z)$ が $[\text{Avg}(x; z)]^2$ にほぼ比例していることを示す。このことは、真の被曝に対して2次のモデルを採用した場合 $\text{Avg}(x; z)$ の2乗を2次項としてモデルを当てはめたとしても、線量反応曲線の形状について良好な推定が得られることを示している。ただし、この2次の係数は、表4に示した数値に1を加えた分だけ偏りが大きくなる。しかし、一般的に非線形モデルの場合については、本報告における線に沿って更に解析を進めることは有用であろう。Pierce及びVaeth¹は、線形2次曲線モデルに関する推論に対して、被曝推定誤差がいかなる影響を及ぼすかを検討し、特にこの種の誤差のために線形からの逸脱がどの程度であるかに注目した。

被曝線量が4 Gyを超える者を解析から除外した場合、被曝推定誤差の影響がどの程度小さくなるかが表6から分かる。しかし、Preston及びPierce¹⁵並びにPierce及びVaeth¹も記述しているように、実際の作業に当たってこの制限を設けるのは、主としてこの理由のためではない。もっと重要な問題として、これ以上の範囲に及ぶ線量域では、真の線量反応が非線形である可能性が存在すること、また、低線量被曝のリスク推定について主として関心があることのためであった。しかし、本報告では、この種の制限を加えることによって、線量が極めて高い場合の誤差モデルの妥当性に対する懸念を減少する効果がある。

線量反応解析には、被曝推定誤差に関係する偏り以外にも問題がある。特に、放射線感受性にどの程度の生物学的個人差があろうかということについて相当の関心もたれている。この問題は発癌リスクの推定に関係してくる可能性がある。すなわち、この種の個人差がかなりあるとすれば、急性影響を耐えて生き残った者は、一般集団の者よりも発癌リスクの点で感受性が低いということが示唆される。この問題を検討するための一つの方法は、簡単にいえば、被曝が同程度の者の中で各種の放射線の影響(癌、急性症状、染色体異常、など)間に正の相関があるか否かを調べることである。この方法については伴ら、¹⁶箱田ら、¹⁷などの報告を参照のこと。ただし、被曝推定値の確率的誤差のために見せかけ上の正の相関がもたらされることは確かである。推定被曝線量が(ほぼ)同じである者の間でも、実は真の被曝線量を示す分布があり、各種の放射線影響のリスクは推定被曝線量に依存するのではなく、この真の被曝線量に依存する。したがって、有効な検討を進めるためには、正の相関が認められても、それが被曝推定誤差に起因する相関よりも大であるか否かを考慮せねばならない。本報告で使用した統計的方法は、この種の計算にも拡張できる。

結論としては、この分野の研究を更に進めることは重要であるが、データ解析のほとんどに本報告で開発した方法の利用を開始することが望ましいと考える。今後行われる改良の中で最も重要なものは、被曝推定誤差モデルの改良であり、DS86線量体系にどの程度の影響が及ぼされるかについて、放影研及び線量の専門家による確認が優先的に行われる必要がある。

その中で、誤差の大きさの遮蔽区分への依存性、及び確率的誤差と系統的誤差との区別についての検討が含まれるべきである。しかし、これは困難な作業であり、ここに示した感度分析によれば、その結果が得られるまでの間は、補正を全く行わないよりは、下記に示す誤差モデルを多くの目的に利用することがはるかによいことが示唆される。

上記に簡単に示したように、偏りを除去するためにその他の統計的な方法もあるが、当該データの場合には、ここに示した被曝線量推定値の補正方法が特に適当であると考えられる。その主な理由は、被曝推定値を多種多様な目的に利用せねばならないからである。すなわち、各種の癌、染色体異常、癌以外の疾病による死亡率、臨床的な各種指標、遺伝的影響及び急性影響についての線形リスクを推定したり、この種の指標に対する線量反応の形状を検討するために線形2次曲線モデルを当てはめたりするときに利用されねばならないからである。

可能なとき、特に線量反応の形状を解析するような特定の目的の場合は、誤差モデルの選択に特別な注意を払い、一層の感度分析を行うことが妥当である。しかし、多くの報告では、多数の課題を取り上げているために既に複雑な内容となっているので、誤差モデルに特別な注意を向けることは不可能である。この種の調査研究では、線形線量反応モデルを通常使用しており、ここに示した感度分析によれば、いかなる誤差モデルを選択するか起因するパラメーター推定値の不確実性は、その他の不確定要素に比べて小さいことが示唆された。この種の検討では、ある与えられた簡単な誤差モデルを選択し、補正の影響を示すようにすることを主体にすれば最善であるかもしれない。誤差の補正は近似値であり、有効な解析を行うために利用されているほかの近似、例えば、線形リスク推定値の使用と同種のものであることを理解せねばならない。

以上の点を考慮し、この誤差の補正という問題に特別の実質的な考慮を払うことが不可能な状況下で解析を行う場合に、利用する誤差モデルを提案することは有用であると考えられる。上記の結果及び考察、又は以前の解析やもっと基本的な統計学的考察に基づき、対数正規モデルをこの目的に利用すべきであることが示唆される。著者の意見では、30%誤差モデルの選択はやや楽観的であるように思われる。このように考える主な理由は、Jablon²の判断として「本節に述べた誤差源による標準誤差は、両市とも合計約±30%になる」との記述と、コメントとして「(特定の面に関する)推定方法にかんがみて、われわれが過小評価しているように思わないこともない」と述べていることにある。Jablonが取り上げた不確定要素は、大部分、被曝者のいた場所と遮蔽状況の評価の不正確性に基づくものであり、それは新旧両線量体系において共通の人力事項である。DS86線量体系におけるほぼ30%弱の誤差⁶が現在の検討対象になっているが、被曝者のいた場所及び遮蔽状況の入力パラメーターを額面通りに受け取った場合、この誤差のある程度は、線量体系それ自体に起因する技術的誤差に関係すると思われる。

以上に基づき、差し当たっては、対数正規35%モデルに若干の特別な焦点を置くことを著者は提案する。その数式は表5に示したが、このモデルの下での被曝の補正に用いる数値

を表3に掲載しなかったもので、ここに示す。

推定値	0.50	1.00	2.00	3.00	4.00	5.00	6.00
補正值							
広島	0.50	0.95	1.79	2.56	3.28	3.98	4.64
長崎	0.50	0.98	1.86	2.68	3.46	4.20	4.92

このモデルについて、表4に相当する数値は0.11～0.12の範囲にわたる。

付 録

線量反応解析の細部についてここで若干の追加説明を行う。

癌データについては、相対的リスクが線量に対して線形であることを前提とする比例危険モデルのみを考慮する。真の被曝 x の人の2値観察値 y を次のとおり表すことができる。

$$y = S[\alpha + \beta x] + \text{誤差} \quad (1)$$

ここで S は、ある対象者についての観察期間中における基礎危険の積分で、確率変数である。値 S は、推定しようとしているパラメーターを含む場合があり、層化に関係することが多い。観察期間は任意であるので、(1)における「誤差」の平均及び分散は、数量 $[y - S(\alpha + \beta x)]$ によって定義される。その平均はゼロであり、分散は S の期待値を $[\alpha + \beta x]$ で掛けたものになる。(1)について、分散を少し修正して $px = S[\alpha + \beta x]$ として加重最小自乗値を反復計算すると、それはこのモデルの最大尤度推定値に等しい。これは、 S が一定で、 y がポアソン分布を示し、かつ、その平均が $S[\alpha + \beta x]$ である場合にも起こるので、これを「ポアソン回帰分析」と言う。

染色体異常データについては、2項モデルの場合、真の被曝線量 x を有する対象者の観察値を次のように示すことができる。

$$y = \alpha + \beta x + \text{誤差} \quad (2)$$

ここで $\text{Var}(\text{誤差}) = px(1 - px)/n$ 。ただし、 $px = \alpha + \beta x$ であり、 n は各対象者について検査した細胞数である。この場合も、反復加重最小自乗値は最大尤度推定値に等しい。

被曝線量推定値に誤差がある場合、推定被曝線量 z のコホートから無作為抽出した被曝者のデータは、上記の二つの場合について下記のように表すことができる。

$$y = S[\alpha + \beta \text{Avg}(x:z)] + \text{誤差}^* \quad (3)$$

$$y = \alpha + \beta \text{Avg}(x:z) + \text{誤差}^* \quad (4)$$

ここで誤差*は、(1)と(2)における誤差、及び分布 $f(x:z)$ を有する確率変数 $[x - \text{Avg}(x:z)]$ に関係した数量の双方を含む。(3)の場合、Sは癌発生までの期間に依存するため、確率変数 Sと被曝推定誤差との間の依存性は極めて小さいので無視する。これについては、Prentice^{4,5}を参照。この近似式により、標本抽出変動とzを与えられたときのxの分布の双方について平均化すれば、この二つの場合はともに誤差*の平均はゼロである。したがって、この意味では最小自乗推定値は不偏である。単純ではあるが、やや長い計算から次が得られた。

$$\text{Var}(\text{誤差}^*) = \begin{cases} pz & \text{[(3)の場合]} \\ pz(1-pz)/n + \beta^2[(n-1)/n] \text{Var}(x:z) & \text{[(4)の場合]} \end{cases}$$

ここで、前記の px の表記方法のやや乱暴な応用であるが、この二つの場合についてそれぞれ次のように書き改めた。 $pz = S[\alpha + \beta \text{Avg}(x:z)]$ 及び $[\alpha + \beta \text{Avg}(x:z)]$ 。癌データの場合、 pz は厳密な意味で誤差*の分散ではなく、前記(1)と同様、Sの期待値の代わりにその観察値を代入している。ここで重要な点は、(3)における回帰の加重値が、(1)の標準的な方法で使われるxの関数と同様に $\text{Avg}(x:z)$ の関数であることである。したがって、zの代わりに $\text{Avg}(x:z)$ を代入すれば、通常の方法が適用できる。これはかなり予想外のことであり、未知のxの代わりに $\text{Avg}(x:z)$ を使用しても、(3)の変動が大きくなることを示している。この結果を導くための計算は、Sが確率的であることに左右されるが、他方、Sを固定した場合に発生する見かけ上の過剰分散が pz よりもはるかに小さくなる。また、この計算は、Sが分布 $f(x:z)$ におけるxとは独立であるとの近似を設けたことにも依存する。しかし、詳細に調べると、今回の状況の下では、このような近似を設けたことの影響は極めて小さいと認められた。

2項モデルの場合、最初の項は、標準的な方法に対する修正として単にzの代わりに $\text{Avg}(x:z)$ で置き換えた分散を示す。第2の項は、考慮しなければならない過剰分散を示す。2項分散の $pz(1-pz)/n$ における係数nのため、また、 pz が大きいこともあり得るので、この過剰分散の項は2項分散とほぼ同じ程度であり、したがって、無視できない。本文で示したように、与えられた誤差モデルに対し、 $CV^2(x:z)$ はzに関してほぼ一定であるので、その過剰分散の項は $\text{Avg}(x:z)^2$ にほぼ比例する。放影研ではかなり以前から次のモデルを用いて過剰分散を補正してきた。

$$\text{Var}(\text{誤差}) = px(1-px)/n + \theta px^2$$

ここで θ はデータから推定されている。例えば、Prestonら¹⁴を参照。しかし、 α は極めて小さいので、これは過剰分散が $\text{Var}(x:z)$ に比例する前記モデルを使用する場合と実質的に同じである。zの代わりに $\text{Avg}(x:z)$ を代入して現行の方法を利用することは、上記で提案した過剰分散を含めた解析と本質的な差はない。ただ、過剰分散の程度 θ は、 $\text{Var}(x:z)$ に関する式から求めるのではなく、データから推定する。この方法は妥当であり、実はこれが

むしろ望ましい。すなわち、この種のデータでは、被曝線量誤差以外にも過剰分散の原因があると考えることは自然であるからである。

以上のように、誤差を伴わない被曝線量を用いるならば、この反復加重最小自乗法は最大尤度法に等しい。誤差がある場合には、このことは当てはまらない。すなわち、最大尤度法は極めて複雑な計算を必要とし、日常的な利用には不適當である。その計算の中には、分布 $f(x:z)$ に関して、 x についての尤度関数の数値積分が含まれる。しかし、前記のもっと簡単な加重最小自乗法との比較を試みることに興味がある。放影研におけるデータの状況に類似するデータを用いて、最大尤度法及び加重最小自乗法の双方を活用して、シミュレーションを行った。一般的な理論からも予想されたように、両者の間で結果にほとんど差はなかった。一般的には、ここに提案した加重最小自乗法の代わりに最大尤度法を使用する現実的な理由はないように思う。

参考文献

1. Pierce DA, Vaeth M: The shape of the cancer mortality dose-response curve for the A-bomb survivor data. (RERF TR in preparation)
2. Jablon S: Atomic bomb radiation dose estimate at ABCC. ABCC TR 23-71
3. Gilbert ES: Some effects of random dose measurement errors on the analyses of atomic bomb survivor data. *Radiat Res* 98:591-605, 1984 (RERF TR 12-82)
4. Prentice RL: Covariate measurement error and parameter estimation in Cox's regression model. *Biometrika* 69:331-42, 1982
5. Prentice RL: Covariate measurement errors in the analysis of cohort and case-control studies. In *Survival Analysis*, eds Crowley J and Johnson R, Institute of Mathematical Statistics Monograph Series, Vol 2, 1982
6. Roesch WC (ed): US-Japan Joint Reassessment of Atomic Bomb Radiation Dosimetry in Hiroshima and Nagasaki. Final Report, Volumes 1 and 2. Radiation Effects Research Foundation, Hiroshima, 1987
7. Cochran WG: Errors of measurement in statistics. *Technometrics* 10:637-66, 1968
8. Proceedings of Workshop on Errors-in-Variabiles in Epidemiology, National Cancer Institute. (To appear in *Statistics in Medicine*, 1988)
9. Fuller WA: *Measurement Error Models*. New York, John Wiley and Sons, 1987
10. Armstrong B: Measurement error in the generalized linear model. *Comm in Statist* 14:529-44, 1985
11. Schafer DW: Covariate measurement errors in generalized linear models. *Biometrika* 74:385-9, 1987

12. Prentice RL: Binary regression using an extended beta-binomial distribution, with discussion of correlation induced by covariate measurement errors. *J Am Statist Assn* 81:321-7, 1986
13. Clayton DG: Models for the analysis of cohort and case-control studies with inaccurately measured exposures. In *Statistical Models for Longitudinal Studies of Health*, ed Dwyer JH. Oxford University Press, 1988
14. Preston DL, McConney ME, Awa AA, Ohtaki K, Itoh M, Honda T: Comparison of dose-response relationships for chromosome aberration frequencies between the T65D and DS86 dosimetries. RERF TR 7-88
15. Preston DL, Pierce DA: The effect of changes in dosimetry on cancer mortality risks estimates in the atomic bomb survivors. *Radiat Res* 114:437-66, 1988 (RERF TR 9-87)
16. Ban S, Awa AA, Ezaki H, Hiraoka T, Yamane M, Nishiki M, Dohi K, Miller RC, Bender MA, Setlow RB: Radiosensitivity and chromosomal aberrations of human skin fibroblast cell strains from atomic-bomb survivors with breast cancer. (RERF TR in preparation).
17. Hakoda M, Akiyama M, Hirai Y, Kyoizumi S, Awa AA: In vivo mutant T cell frequency in atomic bomb survivors carrying outlying values of chromosome aberration frequency. *Mutat Res* 202:203-8, 1988 (RERF TR 19-88)