

---

**Technical Report Series**

---

# **The Effect of Diagnostic Misclassification on Noncancer and Cancer Mortality Dose Response in the RERF Life Span Study**

**Richard Sposto, Ph.D.; Dale L. Preston, Ph.D.;  
Yukiko Shimizu, D.M.Sc.; Kiyohiko Mabuchi, M.D.**



**Radiation Effects Research Foundation**

**A Cooperative Japan-United States Research Organization**

### **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, in part or in toto, as the basis for publication in the open scientific literature. Although they may be quoted and cited, these reports are considered to be internal publications of the Foundation. Copies are available upon request from Publication and Documentation Center, RERF, 5-2 Hijiyama Park, Minami-ku, Hiroshima, 732 Japan.

*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.*

## Errata

Sposto R, Preston DL, Shimizu Y, Mabuchi K: The effect of diagnostic misclassification on noncancer and cancer mortality dose response in the RERF Life Span Study. RERF TR 4-91.

Please note the following corrections.

Page 9, Table 4:

First column, second row:

Null vs. L  $\chi_1^2$  ( $p$  value)

First column, third row:

L vs. L-P  $\chi_1^2$  ( $p$  value)

First column, sixth row:

Percent change in ERR (vs. no correction)

死亡診断書による死因の誤分類が放影研寿命調査における癌および癌以外の死亡率の線量反応に及ぼす影響<sup>§</sup>

## The Effect of Diagnostic Misclassification on Noncancer and Cancer Mortality Dose Response in the RERF Life Span Study

Richard Sposto<sup>1</sup> Dale L. Preston<sup>1</sup> 清水由紀子<sup>2</sup> 馬淵清彦<sup>2,3</sup>

### 要約

放射線被曝による癌以外の死亡のリスク増加<sup>1</sup>が、単に死亡診断書における死因の誤分類に起因するものかどうかを明らかにするために、放影研寿命調査における癌および癌以外の死亡率について解析を行った。癌および癌以外の誤分類率とその死亡年齢への依存性は、1961年から1975年まで実施された寿命調査対象者の一連の剖検から推定した。癌が癌以外の死因と誤分類される割合は20%、癌以外の死因が癌と誤分類される割合は2.8%であった。癌以外の死亡率の線量反応は依然有意であったが、誤分類を補正すると、対象集団における被曝線量1Gyの推定相対過剰リスクは21%、過剰死亡数は23%減少した。一方、癌の被曝線量1Gyの推定相対過剰リスクは12%、過剰死亡数は16%増加した。癌以外の死亡率の線量反応の統計的有意性は、癌の誤分類率を過小評価してもさほど影響を受けなかったが、癌の誤分類が線量と正の関係にあると仮定した場合、統計的有意性は線量増加と共に低下した。誤分類の補正におけるEMアルゴリズムの適用、および複数の死因にその方法を拡大適用することについて検討する。

<sup>§</sup>本業績報告書は研究計画書RP1-75に基づく。本報告の日本語版は別に発行する。承認 1991年7月22日、印刷1992年10月。

放影研<sup>1</sup>統計部、<sup>2</sup>疫学部、および<sup>3</sup>病理疫学部。

---

**Technical Report Series**

---

# The Effect of Diagnostic Misclassification on Noncancer and Cancer Mortality Dose Response in the RERF Life Span Study<sup>§</sup>

Richard Sposto, Ph.D.<sup>1</sup>; Dale L. Preston, Ph.D.<sup>1</sup>;  
Yukiko Shimizu, D.M.Sc.<sup>2</sup>; Kiyohiko Mabuchi, M.D.<sup>2,3</sup>

## Summary

We performed analyses of cancer and noncancer mortality in the Radiation Effects Research Foundation (RERF) Life Span Study (LSS) to determine whether the observed increased risk of noncancer death due to radiation exposure<sup>1</sup> could be attributed solely to misclassification of causes of death on death certificates. Cancer and noncancer misclassification rates and their dependence on age at death were estimated from a series of autopsies performed on LSS participants between 1961 and 1975. The crude misclassification rates were 20% for cancer and 2.8% for noncancer. Although the noncancer dose response remained significant, correcting for this amount of misclassification reduced the estimated noncancer excess relative risk (ERR) at 1 Gy exposure by 21% and the number of excess noncancer deaths in the cohort by 23%. The estimated cancer ERR at 1 Gy was increased by 12% and the excess cancer deaths by 16% as a result of the correction. The statistical significance of the noncancer dose response was relatively insensitive to underestimating the cancer misclassification rate, but was sensitive to assuming that cancer misclassification was positively associated with dose. We discuss the implementation of the EM algorithm for adjusting for misclassification and extensions of the method to more than two causes of death.

## Introduction

The Radiation Effects Research Foundation (RERF) Life Span Study (LSS)<sup>2</sup> is a large cohort of survivors of the atomic bombings of Hiroshima and Nagasaki. Defined in the late 1950s under the auspices of RERF's predecessor, the Atomic

---

<sup>§</sup>This technical report is based on Research Protocol 1-75. The complete text of this report will be available separately in Japanese; approved 22 July 1991; printed October 1992.

Departments of <sup>1</sup>Statistics, <sup>2</sup>Epidemiology, and <sup>3</sup>Epidemiologic Pathology, RERF.



Bomb Casualty Commission (ABCC), with follow-up since October 1950, and expanded several times, the cohort is still being followed today. A major purpose of the LSS is to assess the effect of radiation exposure on mortality. Analyses of mortality in this cohort reveal an increased risk of death among radiation-exposed persons from many types of cancer.<sup>3</sup> One also observes a significant dose response for mortality from causes other than cancer, with the dose effect appearing primarily at high doses.<sup>1</sup> It has been suggested that the observed dose response in noncancer mortality is the result of the misclassification of cancer deaths as noncancer on death certificates. This question is important because if the observed noncancer dose response is due to misclassified cancer cases, then cancer mortality risk estimates obtained from current analyses may be underestimated. An equally important issue, however, is whether or not the RERF data on noncancer mortality are consistent with other epidemiologic and laboratory data that suggest that the primary effect of radiation on human mortality is through increased risk of cancer.<sup>4</sup>

One can approach the problem of misclassification by verifying the current death-certificate-based diagnoses against independent corroborating information. Unfortunately, autopsy results are available for only about 6,500 of the more than 40,000 deaths that occurred among LSS members through 1985. The majority of autopsies were performed in the 1960s and early 1970s and thus apply largely to persons exposed to radiation late in life. One can also obtain indirect information about diagnostic misclassification by looking at the death certificate diagnoses of individuals listed in the RERF tumor registry as having one of a number of relatively fatal cancers or by examining the medical records of persons who had been followed as part of the RERF Adult Health Study (AHS),<sup>5</sup> but these sources also are available for only a fraction of the deaths. Attempts are currently underway to utilize these resources. However, a statistical analysis investigating whether the observed noncancer mortality dose response could reasonably be attributed to misclassification of cancer is a relatively simple first step that can provide much useful information.

In this report we describe such statistical analyses. Functions describing the probabilities of misclassification of cancers and noncancers are estimated from data on 3,508 autopsies conducted on LSS participants from the beginning of 1961 to the end of 1975, the most active period of the intensive RERF autopsy procurement program.<sup>6</sup> These functions are used to correct the numbers of cancer and noncancer deaths for misclassification, using the EM algorithm,<sup>7</sup> in conjunction with Poisson regression methods for grouped data<sup>8</sup> to obtain maximum likelihood estimates of the parameters in the cancer and noncancer mortality functions.

The statistical methods used are closely related to those developed for the problem of uncertainty of cause of death in carcinogenicity experiments. Dinse<sup>9</sup> and Kodell and Chen<sup>10</sup> describe methods for estimating tumor prevalence or mortality when there is a third category of uncertain causes of death in addition to deaths definitely attributed to tumor or definitely attributed not to tumor. Dinse<sup>11</sup> extends this to estimation of tumor incidence. Archer and Ryan<sup>12</sup> take a slightly different approach for hypothesis testing by assuming two cause-of-death categories but allowing cause of death to be misclassified. Another related method presented by Espeland et al.<sup>13</sup> uses the EM algorithm in a Poisson

analysis of longitudinal data to account for both uncertain incidence time and errors in diagnosis. There has also been much interest, particularly in the epidemiologic literature, in misclassification in two-way classifications. Duffy et al.,<sup>14</sup> Elton and Duffy,<sup>15</sup> Flegal et al.,<sup>16</sup> Greenland,<sup>17,18</sup> Greenland and Robins,<sup>19</sup> Hsieh and Walter,<sup>20</sup> and Savitz and Baron<sup>21</sup> all consider various aspects of the problem of misclassification of covariates. Green<sup>22</sup> and White<sup>23</sup> consider the specific problem of misclassification of disease status and write down basic formulas in the  $2 \times 2$  case that are analogous to those forming the basis of our analysis. Chen<sup>24</sup> and Walter and Irwig<sup>25</sup> consider this problem in a more general context. The problem we consider is different from that in carcinogenicity experiments since we are interested in the effect of misclassification on the dose response for both causes of death. In addition, we use nonlinear models to adjust for modifiers of the baseline mortality rates and the dose effect and to describe the dose effect on the misclassification probabilities. Our problem is somewhat simpler than that faced by Espeland et al.<sup>13</sup> in that misclassification applies only to the cause of death; the time of death is known exactly.

## Materials and Methods

### *Adjusting for misclassification of cancer and noncancer diagnoses*

The method used to adjust for the misclassification of cancer and noncancer is a formalization of the following simple procedure. Let  $D_C$  and  $D_N$  represent the numbers of true cancer and noncancer deaths, and  $d_C$  and  $d_N$  the numbers of cancer and noncancer deaths observed from death certificate data, some of which have been misclassified. Let  $\theta_{CN}$  and  $\theta_{NC}$  be the proportions of true cancers misclassified as noncancer on the death certificate and noncancers misclassified as cancer, respectively. Then the formulas

$$\begin{aligned} \text{avg}(d_C) &= D_C(1 - \theta_{CN}) + D_N\theta_{NC} \\ \text{avg}(d_N) &= D_C\theta_{CN} + D_N(1 - \theta_{NC}) \end{aligned} \quad (1)$$

describe the relationship among  $d_C$ ,  $d_N$ ,  $D_C$ , and  $D_N$ , where "avg( $\cdot$ )" denotes average (expected value). For example, the first equation in (1) states that the average observed cancer deaths comprise true cancer deaths that were correctly classified and true noncancer deaths that were misclassified as cancer. The bottom equation has a similar interpretation for average observed noncancer deaths. We can observe  $d_C$  and  $d_N$  directly and use these as estimates of  $\text{avg}(d_C)$  and  $\text{avg}(d_N)$ . Therefore, if we obtain values for  $\theta_{CN}$  and  $\theta_{NC}$ , we can solve for  $D_C$  and  $D_N$  and base our analysis of cancer and noncancer mortality on these corrected quantities.

In the formalization of this method, which is described in detail in the Appendix, the correction is based on the expected values of  $d_C$ ,  $d_N$ ,  $D_C$ , and  $D_N$ , which, along with  $\theta_{CN}$  and  $\theta_{NC}$ , are functions of other variables, such as age, time, age at exposure, city, sex, and radiation dose. Thus, although the concept is essentially the same, the correction for misclassification used in the analysis is more complicated than described above. The formal method allows one to compute likelihood ratio tests and to obtain maximum likelihood estimates of the

parameters in the models that describe the cancer and noncancer mortality function.

### **The DS86 cohort**

The cohort used in these analyses comprises 86,520 members of the LSS with known follow-up between October 1, 1950, and December 31, 1985, for whom DS86 doses<sup>26</sup> have been computed. This cohort contains approximately 10,000 more people than were included in the original DS86 cohort.<sup>27</sup> The increase is a result of extensions to the DS86 system that were made in 1989. These extensions, which are called the Version 3 dosimetry, are described by Fujita.<sup>28</sup> These data were summarized in a 3,640-cell table of person-years and death counts grouped by city of exposure (Hiroshima or Nagasaki); sex; age at the time of exposure (13 categories); absorbed radiation dose to the intestine, in gray (10 categories); and time since exposure (7 categories). Estimated kerma doses greater than 6 Gy were set equal to 6 Gy. Intestinal doses were selected as a reasonable representative organ dose for the many disease types and sites considered. Neutron and gamma components of dose were weighted with a constant RBE of 1. The sample comprises 62% females and 67% persons exposed in Hiroshima. Tables 1 and 2 summarize other characteristics of this cohort.

### **RERF autopsy data**

Data obtained by the ABCC autopsy program between January 1961 and December 1975<sup>29</sup> were used to obtain models for the misclassification probabilities  $\theta_{CN}$  and  $\theta_{NC}$ . Although some autopsy data are available before 1961, a formal program with a goal of obtaining autopsies on all deaths in the LSS cohort was not instituted until 1961.<sup>6</sup> Although the autopsy program continued until the 1980s, the success rate for obtaining autopsies declined rapidly after 1975. Hence, we took 1975 as the closing date for the purposes of this report. Among deaths from individuals in the DS86 cohort, 26% of deaths occurring between

**Table 1.** Summary of sample information

Dose (Gy)	Number of persons	Person-years (10,000s)	Mean dose (Gy)	No. of cancer deaths	No. of noncancer deaths
0.0	36,437	104.63	0.000	2600	10,201
0.01–0.05	27,054	78.00	0.0180	1949	7451
0.06–0.09	5305	15.27	0.0720	411	1509
0.10–0.19	6026	17.34	0.137	482	1701
0.20–0.49	6379	18.35	0.324	543	1785
0.50–0.99	3162	9.09	0.693	298	826
1.00–1.99	1552	4.53	1.35	170	369
2.00–2.99	371	1.04	2.35	40	86
3.00–3.99	149	0.406	3.52	19	51
≥4.0	85	0.237	4.43	10	23
Total/mean	86,520	248.9	0.113	6522	24,002



**Table 2.** Number and types of deaths in the Dosimetry System 1986 cohort (based on death certificates) and the number of these deaths autopsied, 1961–75, by age at death

Age at death (yr)	No. of cancer deaths, 1950–85	No. of autopsied cancer deaths, 1961–75	No. of noncancer deaths, 1950–85	No. of autopsied noncancer deaths, 1961–75
5–9	0	—	0	—
10–19	0	—	83	—
20–29	18	2	353	15
30–39	117	20	525	60
40–49	413	57	918	87
50–59	977	107	1801	144
60–69	1694	293	4045	472
70–79	2485	424	10,185	1127
80–89	796	81	5632	590
≥90	22	1	460	28
Total	6522	985	24,002	2523

1961 and 1975 were autopsied, as compared with about 6.5% for deaths occurring before and after this period. Of the 4,341 autopsies performed in the DS86 cohort after the initiation in 1961 of the formal autopsy program, 3,703 (85%) were performed between 1961 and 1975.

For the analysis of the autopsy data, the underlying cause of death recorded on the death certificate, which is defined as the disease or condition that started the sequence of events leading to death, was compared with the diagnosis designated by the autopsy pathologist as the underlying cause of death.<sup>29</sup> The autopsy data show that most misclassification of cause of death on death certificates occurred between all cancers except leukemia and noncancer disease other than accident. For example, of 1,173 deaths diagnosed as cancer on autopsy, one had been coded as leukemia on the death certificate, whereas 221 had a noncancer death certificate diagnosis. Of 2,355 deaths diagnosed as noncancer on autopsy, 5 were coded as leukemia, 21 as accidents, and 65 as cancer. Since our primary concern is the effect of cancer cases misclassified as noncancer, and since leukemia misclassification as noncancer is rare compared to other cancers misclassified as noncancer, we will assume that deaths due to leukemia and accidents are correctly classified and restrict our attention to misclassification between all cancer except leukemia and noncancer other than accident. Henceforth, the term “cancer” will refer to all deaths due to neoplastic disease with the exception of leukemia, and “noncancer” will refer to all deaths due to causes other than neoplastic disease or accident. A more complete approach, which is outlined in the Appendix, would consider all possible causes of death and types of misclassification between them. However, restricting to these two broad groups of cause of death adequately addresses whether the observed noncancer dose response is due to misclassified cancer, while simplifying the problem considerably. The number of autopsies performed from 1961 to 1975 on cancer and

noncancer deaths in the DS86 cohort are shown in Table 2 by age and death-certificate cause of death.

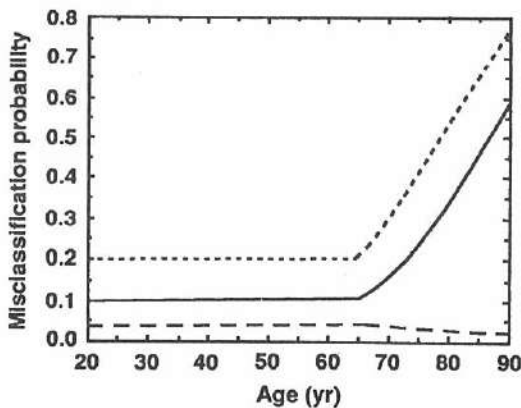
The selection of individuals for autopsy strongly depended on city, sex, time, and radiation dose. There was no evidence in our analysis that selection for autopsy depended on death-certificate diagnosis, although recent, more-comprehensive analyses of misclassification appear to show such an effect. This will be discussed further below.

### **Estimates of cancer and noncancer misclassification rates**

The analyses of the proportion of autopsy-based cancers and noncancers that were misclassified on death certificates were limited to persons for whom DS86 doses were available. Logistic regression methods, using the program GEMBO,<sup>30</sup> were used. The analysis revealed that, at this level of detail in cause of death, the misclassification rates did not depend significantly on sex, city, or time of autopsy. In addition, radiation dose was not a significant modifier of the misclassification probabilities. However, it should be noted that the point estimate for a dose effect on cancer misclassification was relatively large, albeit with a large standard error. Both cancer and noncancer misclassification rates exhibited statistically significant rapid changes in the misclassification rate with age at death for ages over about 65 years. For both types of misclassification a linear logistic model that was constant for age at death less than 65, but depended on age for persons dying after 65 years, captured the essential characteristics of the data. The form of this model is

$$\ln\left(\frac{\theta}{1-\theta}\right) = \gamma_0 + \gamma_1(\text{age} - 65)^+ . \quad (2)$$

The quantity  $(x)^+$  is zero for  $x < 0$  and  $x$  for  $x \geq 0$ , and  $\theta$  is either  $\theta_{\text{NC}}$  or  $\theta_{\text{CN}}$ . The parameter values estimated from the autopsy data were  $\gamma_0 = -2.19$  and  $\gamma_1 = 0.100$  for  $\theta_{\text{CN}}$ , and  $\gamma_0 = -3.25$  and  $\gamma_1 = -0.040$  for  $\theta_{\text{NC}}$ . When averaged over the distribution of age at death, these lead to overall crude misclassification rates of 20% for cancer and 2.8% for noncancer in this sample. These misclassification models were used in the main analysis. The models are illustrated in Figure 1.



Two additional misclassification models were used for sensitivity analyses. First, as a check of sensi-

**Figure 1.** Probability of misclassification as a function of age at death (Equation 2). Solid line:  $\theta_{\text{CN}}$ , with  $\gamma_0 = -2.19$  and  $\gamma_1 = 0.100$  (20% overall cancer misclassification); long dashed line:  $\theta_{\text{NC}}$ , with  $\gamma_0 = -3.25$  and  $\gamma_1 = -0.040$  (2.8% overall noncancer misclassification); dotted line:  $\theta_{\text{CN}}$ , with  $\gamma_0 = -1.374$  and  $\gamma_1 = 0.100$  (36% overall cancer misclassification).

tivity of our results to underestimating the overall cancer misclassification rate, we also performed analyses using a cancer misclassification function with  $\gamma_0$  increased to  $-1.374$  and  $\gamma_1 = 0.100$ , which exactly doubles the rate of misclassification for individuals dying before age 65. For this model, the crude cancer misclassification rate is 36%. This function is also shown in Figure 1. Second, in order to examine the effect of dose dependence on  $\theta_{CN}$ , we considered an extension of (2) that included an additional linear term in dose, with parameter  $\gamma_2$ . The point estimate for this parameter in the linear logistic analysis was 0.23, with a 95% confidence interval of  $(-0.15, 0.56)$ . In the model of  $\theta_{CN}$  used in this part of the sensitivity analysis, we used the parameter values  $\gamma_0 = -2.19$ ,  $\gamma_1 = 0.100$ , and  $\gamma_2 = 0.23$ . Table 3 shows the relationship in this case between cancer misclassification rate and dose for age at death (in years) between 0 and 65 and at 75.

For convenience in the presentation of the results that follow, the three functions for cancer misclassification will be referred to as 20% misclassification, 36% misclassification, and 20% misclassification with dose dependence.

### Regression models for cancer and noncancer mortality

Poisson regression methods<sup>8</sup> were used to model cancer and noncancer mortality. Background rates (death rates for individuals exposed to zero dose) were stratified by city, sex, age-at-exposure category, and time-since-exposure category. For cancer mortality, the dose effect was modeled using the additive relative risk model

$$RR_C(dose, city, sex, agex) = 1 + e^{\alpha_1 city + \alpha_2 sex + \alpha_3 (agex - 25)} \beta_1 dose . \quad (3)$$

Here *sex* is a binary indicator of female sex, *dose* is mean intestinal dose in gray within dose category, and *agex* is mean age at the time of exposure within age-at-exposure categories. In preliminary analyses a squared term in *dose* was included to allow for possible changes in the shape of the cancer dose response as a result of correction for misclassification. Since this squared term was small and not statistically significant in any of the analyses, and since its exclusion had no important effect on the results of interest, only results for cancer relative risk linear in dose are presented. The parameter  $\beta_1$  is interpreted as the excess relative risk (ERR, i.e., the relative risk minus one) at 1 Gy for a male in Hiroshima exposed at age 25 years.

**Table 3.** Percentage of cancer misclassification as a function of dose and age at death

Age (yr)	Dose (Gy)				
	0	0.5	1	1.5	2.0
0-65	10.1	11.2	12.4	13.7	15.1
75	23.4	25.5	27.8	30.1	32.6

The additive relative risk model for noncancer mortality is

$$RR_{NC}(dose) = 1 + \beta_1 dose + \beta_2 dose^2 . \quad (4)$$

This model will be referred to as a linear-power model. Preliminary analyses of total noncancer mortality showed that nonlinearity would be an important characteristic of the noncancer dose response. Although linear-quadratic models have traditionally been used in this context, the linear-power model above avoids the possibility of estimating incidence at doses smaller than the background incidence. This latter feature is also shared by threshold models, which are sometimes used to model data with a high degree of upward curvature, where the effect of radiation is assumed to be zero below a fixed threshold dose. The use of the linear-power model eliminates the need to estimate this threshold value. The power model above provides fits to the noncancer mortality data that are indistinguishable by statistical criteria from those provided by either linear-quadratic or linear-threshold models. The parameter  $\beta_1$  is the ERR at 1 Gy. The parameter  $\beta_2$  reflects the degree of curvature. For example,  $\beta_2 = 1$  corresponds to an ERR that is purely linear in dose and  $\beta_2 = 2$  to a purely quadratic relationship.

Preliminary analyses also revealed no evidence of sex, age-at-exposure, or temporal effects on noncancer relative risk. All of the analyses of noncancer mortality were also performed with the same form of dose-response modification based on city, sex, and age at exposure used in the cancer model above. Since the results based on models that included effect modification differed only negligibly from those based on model (4), only results based on model (4) are presented.

## Results

### *Noncancer and cancer mortality in presence of diagnostic misclassification*

The first two columns of Table 4 summarize the analyses of noncancer and cancer mortality with and without correction for misclassification. The top three rows of the table show the likelihood ratio  $\chi^2$  statistic for three different hypothesis tests of the noncancer dose response: (1) null model (i.e., no dose dependence) versus linear-power model, (2) null model versus linear model, and (3) linear versus linear-power model. The remaining rows of Table 4 give the parameter estimates for the cancer and noncancer relative risk function based on a linear model in dose for cancer and a linear-power model for noncancer. The percentage change in ERR and the estimated number and percentage change in excess cases in the DS86 cohort due to radiation exposure are also shown.

Without correction for misclassification, the linear-power dose response for noncancer mortality is significant, with  $\chi^2_2 = 18.7$  ( $p < .001$ ). Much of this  $\chi^2$  can be attributed to the test of linear increase in relative risk,  $\chi^2_1 = 15.6$  ( $p < .001$ ), although some improvement is obtained by allowing nonlinearity. With correction for 20% misclassification, the linear-power dose response is still significant ( $p < .004$ ), and once again much of this is due to the test of linear increase in relative risk. Note, however, that the estimate of the parameter  $\beta_2$ , which affects the degree of nonlinearity in the noncancer dose response, is greater than 1.5 in both of these cases, indicating in a sense that the model that fits these data well

**Table 4.** Results of likelihood ratio tests of the noncancer dose response and estimates of parameters of the relative risk function for noncancer (Equation 4) and cancer (Equation 3) with and without correction for misclassification

Test of noncancer dose response	Sensitivity analysis		
	No correction for misclass.	Correction for 20% misclass.	Correction for 36% misclass. with $\theta_{CN}$ dose effect
Null vs. L-P $\chi^2$ ( <i>p</i> value)	18.7 (<.001)	11.3 (<.004)	5.9 (.052)
Null vs. L $\chi^2$ ( <i>p</i> value)	15.6 (<.001)	8.7 (.003)	3.8 (.051)
L vs. L-P $\chi_1$ ( <i>p</i> value)	3.1 (.078)	2.6(>.10)	2.1(>.10)
	Noncancer model parameter estimates (SEs)		
$\beta_1$	0.0630 (0.0317)	0.0498 (0.0320)	0.0336 (0.0326)
$\beta_2$	1.560 (0.444)	1.631 (0.554)	1.793 (0.803)
Change in ERR (vs. no correction)	—	-21	-47
Excess noncancer deaths (% change)	129	99 (-23)	64 (-50)
	Cancer model parameter estimates (SEs)		
$\alpha_1$	-0.676 (0.350)	-0.712 (0.327)	-0.674 (0.305)
$\alpha_2$	0.710 (0.294)	0.668 (0.267)	0.602 (0.248)
$\alpha_3$	-0.0416 (0.00946)	-0.0441 (0.00870)	-0.0445 (0.00810)
$\beta_1$	0.494 (0.129)	0.553 (0.131)	0.556 (0.121)
Percent change in ERR (vs. no correction)	—	12	13
Excess cancer deaths (% change)	274	317 (16)	369 (35)
			374 (36)
			62 (-52)
			0.0308 (0.0315)
			1.678 (0.863)
			4.3 (.12)
			2.9 (.089)
			1.4 (>.10)

NOTE: ERR = excess relative risk; L = linear model; L-P = linear-power model; SE = standard error of the estimate.



is closer to a purely quadratic model than to a purely linear model. Correcting for misclassification decreases the estimated noncancer ERR at 1 Gy from 0.0630 to 0.0498, or by 21%. The number of excess noncancer deaths declines from 129 to 99, or by 23%. The cancer ERR increases from 0.494 to 0.553, or by 12% (evaluated for Hiroshima males exposed at age 25 years), and the number of excess cancer deaths from 274 to 317, or by 16%. In this case, correction for misclassification results in a net 832 cases being reclassified from noncancer to cancer, which decreases the number of noncancer deaths by 3.5% and increases the number of cancer deaths by 12.8%.

### **Sensitivity analyses**

The results of the sensitivity analysis, that is, with correction for 36% misclassification and for 20% misclassification with dose dependence, are shown in the last two columns of Table 4. In the first sensitivity analysis we found that the dose response in noncancer mortality was flatter than before and was reduced to borderline significance ( $p = .052$ ) after correcting for 36% misclassification. In this case the estimated noncancer ERR decreases by 47% and the number of excess noncancer deaths by 50%. The estimated cancer ERR increases by 13% but the number of excess cancer deaths by 35%. Although doubling the cancer misclassification rate resulted in only a small increase in cancer ERR, the number of excess cancer deaths increased substantially, because correcting for misclassification increases the estimated background cancer rate. In the second sensitivity analysis, the significance of the noncancer dose response disappears after correcting for 20% misclassification with dose dependence ( $p = .12$ ). The noncancer ERR decreases by 52%, and cancer ERR increases by 37% compared to no correction. The corresponding changes in noncancer and cancer excess deaths are 52% decrease and 36% increase, respectively.

### **Discussion**

An analysis of the noncancer and cancer mortality in the LSS demonstrates that the observed significant dose response in noncancer mortality cannot easily be explained by dose-independent misclassification of cancer as noncancer. Correcting for the amount of misclassification estimated from the RERF autopsy series, 20% cancer and 2.8% noncancer misclassification, reduces the significance level of the noncancer dose response and reduces the estimated ERR by 18% and the number of excess noncancer deaths by 23%. However, the noncancer mortality dose response remains significant ( $p = .004$ ). The correction for misclassification also results in a 12% increase in the estimated ERR for cancer and a 16% increase in the number of excess cancer deaths.

Sensitivity analyses showed that the statistical significance of the noncancer dose response was relatively insensitive to underestimating the overall cancer misclassification rate. The noncancer dose response was still on the border of significance ( $p = .052$ ) after almost doubling the overall cancer misclassification rate to 36%. However, these analyses also showed that the significance of the noncancer dose response was sensitive to whether cancer misclassification was allowed to depend positively on dose. When we assumed a dose dependence corresponding to a roughly 20% increase in the cancer misclassification rate at

---

1 Gy, the noncancer dose response was no longer significant. In this case, however, the estimated cancer ERR at 1 Gy increased by 37%.

Note that this last sensitivity analysis assumes that misclassification of cancer cases systematically occurred more frequently in high-dose groups than in low-dose groups. One interpretation of this would require the physician making the diagnosis of cause of death to have some idea about an individual's exposure level and consciously or unconsciously to use this information to alter diagnostic criteria so that cancers would be more often misdiagnosed as noncancer in high-dose cases. It is unlikely that the people making death certificate diagnoses have any specific information on dose, although they may know an individual's approximate location relative to the hypocenter, which is correlated with dose. This type of dependence of misclassification on dose seems to us counterintuitive, since one might expect that deaths would more likely be attributed to cancer, rather than to noncancer, in the case of higher radiation exposure. Another possibility is that radiogenic cancers also happen to be those that are most difficult to diagnose, so that a dose response in misclassification will be observed if type of cancer is not taken into account. More-detailed analyses of misclassification, which are currently under way, may clarify whether a dose response for misclassification probability should be included in the type of analysis that we have described.

We found from autopsy data that cancer and noncancer misclassification rates did not depend significantly on city, sex, or time since exposure. However, there was a very strong dependence of the misclassification rates on age at death, which was taken into account in our models. The evidence from the autopsy data about dependence of misclassification rates on dose is weak. There was evidence that the selection of patients for autopsy depended on city, sex, time since exposure, dose, and age at exposure but not on the death certificate cause of death after accounting for these other effects, although recent analyses of misclassification may suggest such an effect. The last is important because, if the selection rates depend in an important way on the initial diagnosis, our estimates of the cancer misclassification rates could be biased. We did, however, use sensitivity analysis to investigate the effect of having underestimated the cancer misclassification rate.

Our analysis of the autopsy data was not comprehensive but was carried out to meet the specific objective of this report. A recent, more comprehensive analysis of misclassification points out some areas where adjustments for misclassification can be refined. For example, the estimated cancer misclassification rate appears to depend on place of autopsy (RERF or hospital) and on place of death (hospital or home). In the former case, hospital autopsy results might have been considered when completing the death certificate. Estimates of the overall cancer misclassification rate based only on RERF autopsies are approximately 25% rather than the 20% which we used. This is still well within the range of values considered in our sensitivity analyses. In the latter case, deaths occurring in hospitals are more accurately diagnosed than those occurring at home. In future analyses it may be important to make this distinction, since the distribution of home versus hospital deaths in the autopsy series may be different than in the general population, implying that the "average" misclassification rates we are using are not appropriate.

In this report we have treated the estimated misclassification functions as fixed and made inferences about cancer and noncancer mortality, rather than estimating the misclassification functions using the combined mortality and autopsy data. Although it is possible to estimate jointly the mortality function and misclassification rates using the combined data, we have chosen in this report to focus on the effect of misclassification on inference about cancer and noncancer mortality dose response. To do this, we used the autopsy data to estimate reasonable models for the misclassification mechanism that is operating in the LSS, considered these models fixed, and used them as a basis for the correction for misclassification.

To refine the correction for misclassification for a specific cause of death (e.g., heart disease), it is possible simply to create two categories, one for heart disease and the other for all other causes of death, estimate the misclassification probabilities between these two categories, and apply the methods described above. The advantage of this approach is that it is relatively straightforward to apply. The disadvantages are that the broad, catchall category comprising very different causes of death may not be interpretable biologically and that a separate analysis must be performed for every category of death of interest. Alternatively, one could extend the procedure, as outlined in the Appendix, to more than two classifications that group similar causes of death. This is more complicated to implement but may have advantages in interpretability and internal consistency. Either approach would be useful in determining whether a dose-related increase in the number of deaths due to a specific noncancer cause can be explained by misclassification of cancer or to refine estimates of radiation risk for particular types of cancer, such as liver cancer, that are particularly susceptible to diagnostic misclassification on the death certificate.

## References

1. Shimizu Y, Kato H, Schull WJ, Hoel DG: Life Span Study Report 11. Part 3. Noncancer mortality, 1950–1985, based on the revised doses (DS86). RERF TR 2-91
2. Beebe GW, Ishida M: Joint JNIIH–ABCC study of life span of atomic bomb survivors. Research plan. ABCC TR 4-59
3. Shimizu Y, Kato H, Schull WJ: Studies of the mortality of A-bomb survivors. 9. Mortality, 1950–1985: Part 2. Cancer mortality based on the recently revised doses (DS86). *Radiat Res* 121:120–41, 1990
4. National Research Council: *Health Effects of Exposure to Low Levels of Ionizing Radiation* (BEIR V). Washington, D.C., National Academy Press, 1990
5. Hollingsworth JW, Beebe GW: ABCC–JNIIH Adult Health Study: Provisional research plan. ABCC TR 9-60
6. Zeldis LJ, Matsumoto YS: ABCC–JNIIH pathology studies, Hiroshima–Nagasaki, provisional research plan. 1. Description and scope of program. ABCC TR 4-61
7. Dempster AP, Laird NM, Rubin DB: Maximum likelihood from incomplete data via the EM algorithm. *J Royal Stat Soc Ser B* 39:1–38, 1977
8. Clayton D: The analysis of event history data: A review of progress and outstanding problems. *Stat Med* 7:819–41, 1988

9. Dinse GE: Nonparametric prevalence and mortality estimators for animal experiments with incomplete cause-of-death data. *J Am Stat Assoc* 81(394):328–36, 1986
10. Kodell RL, Chen JJ: Handling cause of death in equivocal cases using the EM algorithm. *Commun Stat Theory Methods* 16(9):2565–85, 1987
11. Dinse GE: Estimating tumor incidence rates in animal carcinogenicity experiments. *Biometrics* 44:405–15, 1988
12. Archer LE, Ryan LM: Accounting for misclassification in the cause-of-death test for carcinogenicity. *J Am Stat Assoc* 84(407):787–91, 1989
13. Espeland MA, Platt OS, Gallagher D: Joint estimation of incidence and diagnostic error rates from irregular longitudinal data. *J Am Stat Assoc* 84(408):972–9, 1989
14. Duffy SW, Rohan TE, Day NE: Misclassification in more than one factor in a case-control study: A combination of Mantel-Haenszel and maximum likelihood approaches. *Stat Med* 8(12):1529–36, 1989
15. Elton RA, Duffy SW: Correcting for the effect of misclassification bias in a case-control study using data from two different questionnaires. *Biometrics* 39(3):659–63, 1983
16. Flegal KM, Brownie C, Haas JD: The effects of exposure misclassification on estimates of relative risk. *Am J Epidemiol* 123(4):736–51, 1986
17. Greenland S: Statistical uncertainty due to misclassification: Implications for validation substudies. *J Clin Epidemiol* 41(12):1167–74, 1988
18. Greenland S: On correcting for misclassification in twin studies and other matched-pair studies. *Stat Med* 8(7):825–9, 1989
19. Greenland S, Robins JM: Confounding and misclassification. *Am J Epidemiol* 122(3):495–506, 1985
20. Hsieh CC, Walter SD: The effect of non-differential exposure misclassification on estimates of the attributable and prevented fraction. *Stat Med* 7(10):1073–85, 1988
21. Savitz DA, Baron AE: Estimating and correcting for confounder misclassification. *Am J Epidemiol* 129(5):1062–71, 1989
22. Green MS: Use of predictive value to adjust relative risk estimates biased by misclassification of outcome status. *Am J Epidemiol* 117(1):98–105, 1983
23. White E: The effect of misclassification of disease status in follow-up studies: implications for selecting disease classification criteria. *Am J Epidemiol* 124(5):816–25, 1986
24. Chen TT: A review of methods for misclassified categorical data in epidemiology. *Stat Med* 8:1095–106, 1989
25. Walter SD, Irwig LM: Estimation of test error rates, disease prevalence and relative risk from misclassified data: A review. *J Clin Epidemiol* 41(9):923–37, 1988
26. Roesch WC (ed): *US–Japan Joint Reassessment of Atomic Bomb Radiation Dosimetry in Hiroshima and Nagasaki: Final Report* (Volumes 1 and 2). Hiroshima, Radiation Effects Research Foundation, 1987
27. Preston DL, Pierce DA: The effects of changes in the dosimetry on cancer mortality risk estimates in the atomic bomb survivors. *Radiat Res* 114:437–66, 1988

28. Fujita S: Versions of DS86. RERF Update 1(2):3, 1989 (Hiroshima, Radiation Effects Research Foundation)
29. Yamamoto T, Moriyama IM, Asano M, Guralnick, L: RERF pathology studies, Hiroshima and Nagasaki. Report 4: The autopsy program and Life Span Study, January 1961–December 1975. RERF TR 18-78
30. Preston DL, Lubin JH, Pierce DA: *Epicure: User's Guide*. Seattle, HiroSoft International, 1990



## Appendix

### *Adjusting for misclassification of cancer and noncancer in the context of Poisson regression*

Assume that the mortality function  $\lambda(t)$  depends on a time-varying vector of covariates  $\mathbf{Z}(t)$ , and approximate this mortality function by a piecewise constant failure function taking values  $\lambda_k$ , where  $k = 1, \dots, K$  indexes one of  $K$  cells of a partition of the space of  $\mathbf{Z}(t)$ . Then the approximate likelihood of the data is proportional to

$$\prod_{k=1}^K \lambda_k^{D_k} \exp[-P_k \lambda_k],$$

where  $P_k$  is the total number of person-years spent by all individuals in the  $k$ th cell and  $D_k$  is the number of deaths observed among those in the cell.<sup>8</sup> In practice, the step function  $\lambda_k$  is described by a continuous function  $\lambda(\mathbf{z}_k)$ , where the covariate vector  $\mathbf{Z}(t)$  is discretized into one of  $K$  values  $\mathbf{z}_k$ ,  $k = 1, \dots, K$ . Various nonlinear models of  $\lambda(\mathbf{z}_k)$  can be fit easily using the computer program AMFIT,<sup>30</sup> which is designed specifically for this problem.

Now, consider one of the cells in the partition referred to above, and define the following quantities (to simplify notation, the subscript  $k$  and arguments have been dropped):

- $P$  = number of person-years;
- $D_C$  = number of true cancer deaths;
- $D_N$  = number of true noncancer deaths;
- $d_C$  = number of observed cancer deaths (from death certificates);
- $d_N$  = number of observed noncancer deaths (from death certificates);
- $D$  = total number of deaths,  $D_C + D_N = d_C + d_N$ ;
- $X_C$  = the number of misclassified true cancer deaths;
- $X_N$  = the number of misclassified true noncancer deaths;
- $\lambda_C$  = cancer mortality rate;
- $\lambda_N$  = noncancer mortality rate;
- $\theta_{CN}$  = cancer misclassification rate—the probability that a true cancer is misclassified as a noncancer on the death certificate; and
- $\theta_{NC}$  = noncancer misclassification rate—the probability that a true noncancer is misclassified as a cancer on the death certificate.

Note that  $\theta_{CN}$  and  $\theta_{NC}$  are false-negative and false-positive rates, respectively, for the diagnosis of cancer. In addition, in terminology commonly used in epidemiology,  $(1 - \theta_{CN})$  and  $(1 - \theta_{NC})$  are the cancer and noncancer detection rates, respectively.

	<i>True cancer</i>	<i>True noncancer</i>	
<i>Death certificate cancer diagnosis</i>	$D_C - X_C$ $\frac{\lambda_C}{\lambda_C + \lambda_N} (1 - \theta_{CN})$	$X_N$ $\frac{\lambda_N}{\lambda_C + \lambda_N} \theta_{NC}$	$d_C = D_C - (X_C - X_N)$
<i>Death certificate noncancer diagnosis</i>	$X_C$ $\frac{\lambda_C}{\lambda_C + \lambda_N} \theta_{CN}$	$D_N - X_N$ $\frac{\lambda_N}{\lambda_C + \lambda_N} (1 - \theta_{NC})$	
	$D_C$	$D_N$	$D$

**Figure A1.** Number of deaths ( $D$ ) in each classification category and the multinomial probabilities for the cells in a cross-classification of deaths by true underlying cause of death and death certificate diagnosis. Marginal totals for deaths are also shown.

Figure A1 describes the number of deaths and the multinomial probabilities for each cell in a cross-classification of deaths by true underlying cause and death-certificate diagnosis. Marginal totals for deaths are also shown.

The probability that a death certificate diagnosis of cancer is really a noncancer ( $\theta'_{CN}$ ) and the probability that a death certificate diagnosis of noncancer is really a cancer ( $\theta'_{NC}$ ) are easily seen to be

$$\theta'_{CN} = \frac{\lambda_N \theta_{NC}}{\lambda_C (1 - \theta_{CN}) + \lambda_N \theta_{NC}}$$

$$\theta'_{NC} = \frac{\lambda_C \theta_{CN}}{\lambda_C \theta_{CN} + \lambda_N (1 - \theta_{NC})} \quad (A.1)$$

In general these probabilities will depend on dose and any other variables that affect  $\lambda_C$  and  $\lambda_N$ ,  $\theta_{CN}$  and  $\theta_{NC}$ . The quantities  $(1 - \theta'_{CN})$  and  $(1 - \theta'_{NC})$  are sometimes called cancer and noncancer confirmation rates, respectively.

The complete data likelihood for one cell of the  $K$ -dimensional person-year table in terms of the true (unobservable) death frequencies  $D_C$  and  $D_N$  and misclassified cases  $X_C$  and  $X_N$  is proportional to

$$L(D_C, D_N, X_C, X_N | \lambda_C, \lambda_N, \theta_{CN}, \theta_{NC}) = [(P\lambda_C)^{D_C} (P\lambda_N)^{D_N} e^{-P(\lambda_C + \lambda_N)}] \cdot$$

$$\left[ \theta_{CN}^{X_C} (1 - \theta_{CN})^{D_C - X_C} \theta_{NC}^{X_N} (1 - \theta_{NC})^{D_N - X_N} \right] \quad (A.2)$$

Note that the first term on the right-hand side of (A.2) is proportional to the joint likelihood for two independent Poisson random variables, whereas the second term is proportional to the likelihoods of two independent binomial random variables. This likelihood is maximized by maximizing separately the two Poisson likelihoods for cancer and noncancer mortality and the two binomial likelihoods for each type of misclassification.

When  $X_C$  and  $X_N$ , and therefore  $D_C$  and  $D_N$ , are not observable, the likelihood of the observable data is proportional to

$$L(d_C, d_N | \lambda_C, \lambda_N, \theta_{CN}, \theta_{NC}) = [P[\lambda_C(1 - \theta_{CN}) + \lambda_N\theta_{NC}]]^{d_C} [P[\lambda_C\theta_{CN} + \lambda_N(1 - \theta_{NC})]]^{d_N} e^{-P(\lambda_C + \lambda_N)}, \quad (A.3)$$

where the quantities in square brackets are Poisson rates for observed cancer and noncancer deaths, which are the Poisson rates for true cancer and noncancer deaths convoluted with the misclassification probabilities. The total complete data likelihood and observable data likelihood are the products of terms like (A.2) and (A.3), respectively, over each of the  $K$  cells.

The log likelihood based on (A.3) can be maximized using the EM algorithm.<sup>7</sup> In the E-step,  $D_C$ ,  $D_N$ ,  $X_C$ , and  $X_N$  in (A.2) are replaced with their expected values given the observed data  $d_C$  and  $d_N$ . From Figure A1 one can see that these expectations are

$$\begin{aligned} E(D_C | d_C, d_N) &= d_C(1 - \theta'_{CN}) + d_N\theta'_{NC} \\ E(D_N | d_C, d_N) &= d_C\theta'_{CN} + d_N(1 - \theta'_{NC}) \\ E(X_C | d_C, d_N) &= d_N\theta'_{NC} \\ E(X_N | d_C, d_N) &= d_C\theta'_{CN}. \end{aligned} \quad (A.4)$$

In the M-step, the likelihood (A.2) is maximized using the expected values obtained from (A.4). Likelihood ratio tests are obtained by computing the likelihood based on (A.3), which can be done easily.

As alluded to in the Discussion, one can in principle directly use the autopsy data for a subset of individuals in this maximization procedure. For the portion of the data with autopsy information,  $D_C$ ,  $D_N$ ,  $X_C$ , and  $X_N$  in equation (A.2), are known. Thus the E-step of the EM algorithm is unnecessary. For deaths without autopsy information, the EM algorithm is applied as described. In the analyses here, however, although we estimate  $\theta_{CN}$  or  $\theta_{NC}$  from the autopsy data, we treat them as known functions. In this case, the maximization involves only the Poisson-like terms of equation (A.2).

### **Extensions to multiple causes of death**

Conceptually, these procedures can be extended to the case of multiple causes of death and misclassifications between them, although implementation is more difficult. Consider  $M$  causes of death, and let  $\lambda_i$  be the mortality function for cause  $i$ ,  $i = 1, \dots, M$ . In addition, let  $\theta_{ij}$  be the proportion of individuals with true cause of death  $i$  who are classified on the death certificate as cause  $j$ ,  $i, j = 1, \dots, M$ .

Similarly, define  $D_i$  and  $d_i$  as the number of true and observed deaths of type  $i$  and  $X_{ij}$  as the number of true type  $i$  deaths misclassified as type  $j$ . Then the likelihood corresponding to (A.2) above is proportional to

$$L[(D_i, X_{ij}, i, j = 1, \dots, M)] = \prod_{i=1}^M (P\lambda_i)^{D_i} e^{-P\lambda_i} \prod_{i,j=1}^M \theta_{ij}^{X_{ij}}, \quad (\text{A.5})$$

subject to the conditions that  $X_{i.} = D_i$  and  $\theta_{i.} = 1$ , where the subscript “.” indicates summation over an index. The likelihood corresponding to (A.3) is proportional to

$$L(d_i, i = 1, \dots, M) = \prod_{i=1}^M \left\{ \sum_{j=1}^M \lambda_j \theta_{ji} \right\}^{d_i} e^{-P\lambda_i}, \quad (\text{A.6})$$

where  $d_i = X_{i.}$  The quantities corresponding to equations (A.1) are

$$\theta'_{ij} = \frac{\lambda_j \theta_{ji}}{\sum_{k=1}^M \lambda_k \theta_{ki}} \quad (\text{A.7})$$

and those corresponding to equations (A.4) are

$$E(D_i | d_1, \dots, d_M) + \sum_{j=1}^M d_j \theta'_{ji}$$

$$E(X_{ij} | d_1, \dots, d_M) = d_j \theta'_{ji}. \quad (\text{A.8})$$

With these equations, the EM algorithm proceeds as before.

---

業績報告書シリーズ  
*Technical Report Series*

---

死亡診断書による死因の誤分類が放影研寿命  
調査における癌と癌以外の死亡率の  
線量反応に及ぼす影響

The Effect of Diagnostic Misclassification  
on Noncancer and Cancer Mortality Dose Response  
in the RERF Life Span Study

Richard Sposto Dale L. Preston 清水由紀子 馬淵清彦



## 放影研業績報告書シリーズ RERF 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、その一部また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。報告書は放影研出版・資料センター、〒732 広島市南区比治山公園5-2で入手できる。

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。

## 業績報告書シリーズ

死亡診断書による死因の誤分類が放影研寿命調査における  
癌と癌以外の死亡率の線量反応に及ぼす影響<sup>§</sup>The Effect of Diagnostic Misclassification  
on Noncancer and Cancer Mortality Dose Response  
in the RERF Life Span StudyRichard Sposto<sup>1</sup> Dale L. Preston<sup>1</sup> 清水由紀子<sup>2</sup> 馬淵清彦<sup>2,3</sup>

## 要 約

放射線被曝による癌以外の死亡のリスク増加<sup>1</sup>が単に死亡診断書における死因の誤分類に起因するものかどうかを明らかにするために、放射線影響研究所(放影研)寿命調査(LSS)における癌の死亡率と癌以外の死亡率について解析を行った。癌および癌以外の誤分類率と、その死亡年齢への依存性は、1961年から1975年まで実施された寿命調査対象者の一連の剖検から推定した。癌が癌以外の死因と誤分類される割合は20%、癌以外の死因が癌と誤分類される割合は2.8%であった。誤分類を補正しても癌以外の死亡率の線量反応は、依然有意であったが、対象集団における被曝線量1 Gyの推定相対過剰リスクは21%、過剰死亡数は23%減少した。一方、癌の被曝線量1 Gyの推定相対過剰リスクは12%、過剰死亡数は16%増加した。癌以外の死亡率の線量反応における統計的有意性は、癌の誤分類率をLSSの剖検データから得られる率よりも大きく見積っても、さほど影響を受けなかったが、癌の誤分類が線量と正の関係にあると仮定した場合、統計的有意性は低下した。誤分類を補正するためにEMアルゴリズムを適用すること、および、複数の死因にその方法を拡大適用することについて検討する。

## 緒 言

放射線影響研究所の寿命調査集団<sup>2</sup>は広島・長崎の原爆被爆者の大規模なコホートである。

<sup>§</sup>本業績報告書は研究計画書RP 1-75に基づく。本報告の英語版は別に発行した。本報告に基づく論文はBiometrics 48:605-17, 1992年に掲載された。承認 1991年7月22日。印刷 1993年3月。

放影研<sup>1</sup>統計部<sup>2</sup>疫学部<sup>3</sup>病理疫学部

1950年代後半にABCCのもとで確立されたこの集団は1950年10月から追跡調査され、数回の拡大を経て、現在も依然として追跡調査が続いている。寿命調査の主な目的は死亡率への放射線被曝の影響を評価することである。このコホートの死亡率解析により、様々な種類の癌による死亡リスクが放射線被曝者において増加していることが明らかになった。<sup>3</sup> また、癌以外の原因による死亡率にも有意な線量反応が認められており、主に高線量で線量の影響が認められる。<sup>1</sup> 癌以外の死亡率に認められた線量反応は、癌による死亡を死亡診断書上で癌以外によると誤分類した結果である可能性が考えられる。観察された癌以外の線量反応が癌症例の誤分類によるのであれば、現在の解析で得られた癌死亡率のリスク推定値は過小推定値かもしれないので、この問題は重要である。しかし、これと同等に重要な問題は、ヒト死亡率への放射線の主な影響は癌リスクの増加によると示唆する他の疫学および実験データ<sup>4</sup>と、放影研の癌以外の死亡率に関するデータが一致するか否かである。

この誤分類の問題は、現在の死亡診断書に基づく診断を、裏付けとなる独立した情報と照合することにより検討することができる。遺憾ながら、1985年の寿命調査対象者40,000人以上の死亡例のうち剖検所見が得られているのは6,500人のみである。剖検のほとんどは1960年代と1970年代前半に行われているので、人生の後半で放射線に被曝した人が大部分である。比較的致命的な癌の一つに罹患したことで放影研の腫瘍登録に載った者の死亡診断書の診断を見ることにより、また、放影研成人健康調査(AHS)<sup>5</sup>の一環として追跡調査した者の診断書を検討することにより、診断の誤分類に関する間接的な情報を得ることもできるが、これら情報源も死亡例の一部のものにしか入手できない。現在これらの情報源の活用を試みているところである。しかし、観察された癌以外の死亡率の線量反応が癌の誤分類によるものであるか否かを検討するために行う統計解析は、多くの有用な情報が得られる比較的簡単な第一歩である。

本報ではこのような統計的解析について述べる。放影研の剖検プログラムが最も活発に行われていた1961年初めから1975年終わりまでに剖検を実施した寿命調査対象者3,508例に関するデータから、癌および癌以外の誤分類の確率を表す関数を推定する。<sup>6</sup> これらの関数を用いて誤分類した癌および癌以外の死亡例数を修正し、EMアルゴリズム<sup>7</sup> およびグループ別データのポアソン回帰法<sup>8</sup> により、癌および癌以外の死亡率の関数におけるパラメータの最尤推定値を求める。

用いた統計的方法は、発癌性実験における死因の不確実性の問題を解決するために開発された方法に密接に関連する。Dinse<sup>9</sup> ならびにKodellおよびChen<sup>10</sup> は、確実に腫瘍による死亡、および確実に腫瘍以外による死亡のほかに、死因が不明確な第3のカテゴリーがある場合に、腫瘍の有病率または死亡率を推定する方法を述べている。Dinse<sup>11</sup> はこれを腫瘍発

生率の推定に拡大応用した。Archer および Ryan<sup>12</sup> は仮説のテストに少し異なる方法を用いており、二つの死因カテゴリーを仮定し死因の誤分類も考慮に入れている。これに関連して Espeland ら<sup>13</sup> が提示した方法では、経時的データのポアソン解析において EM アルゴリズムを用い、不明確な発生時期と診断上の誤りの両方を考慮に入れている。また、特に疫学の文献では、2区分への分類における誤分類について大きな関心が持たれている。Duffy ら、<sup>14</sup> Elton および Duffy,<sup>15</sup> Flegal ら,<sup>16</sup> Greenland,<sup>17,18</sup> Greenland および Robins,<sup>19</sup> Hsieh および Walter,<sup>20</sup> Savitz および Baron<sup>21</sup> は共変量の誤分類における問題を様々な点から検討している。Green<sup>22</sup> および White<sup>23</sup> は疾患状態の誤分類の特定の問題を検討し、 $2 \times 2$  の基本式を書いているが、これは我々の解析の基盤となる式に類似している。Chen<sup>24</sup> ならびに Walter および Irwig<sup>25</sup> はより一般的な観点からこの問題を検討している。我々は両方の死因に関する誤分類の線量反応への影響に関心をもっているので、我々の検討する問題は発癌性実験の場合とは異なる。更に我々は、非線形モデルを用いて基準死亡率と線量効果の修飾因子について補正し、誤分類確率へ及ぼす線量の影響について述べる。死亡時期が判明しており誤分類は死因のみに当てはまるので、我々の問題は Espeland ら<sup>13</sup> が扱った問題よりも幾分単純である。

## 材料および方法

癌および癌以外の診断の誤分類に対する補正 癌および癌以外の誤分類について補正するために用いた方法は、以下の単純な手順を公式化したものである。 $D_C$  と  $D_N$  は真の癌および癌以外の死亡数を表し、 $d_C$  と  $d_N$  は死亡診断書に認められた癌と癌以外の死亡数を表し、それらのうち幾つかは誤って分類されているとする。 $\theta_{CN}$  は死亡診断書では癌以外と誤分類された真の癌症例の割合(癌の誤分類率)、 $\theta_{NC}$  は癌に誤分類された癌以外の症例の割合(癌以外の誤分類率)とする。この場合、式

$$\begin{aligned} \text{avg}(d_C) &= D_C(1 - \theta_{CN}) + D_N\theta_{NC} \\ \text{avg}(d_N) &= D_C\theta_{CN} + D_N(1 - \theta_{NC}) \end{aligned} \quad (1)$$

は  $d_C$ ,  $d_N$ ,  $D_C$  および  $D_N$  の関係を表し、“avg( $\cdot$ )” は平均(期待値)を示す。たとえば、(1)の初めの式は、平均観察癌死亡数が、正確に分類された真の癌死亡数と、癌と誤分類された癌以外の死亡数から成ることを示す。二番目の式は癌以外の平均観察死亡数も同様であることを示している。我々は  $d_C$  および  $d_N$  を直接観察し、これらを  $\text{avg}(d_C)$  と  $\text{avg}(d_N)$  の推定値として用いることができる。したがって、 $\theta_{CN}$  と  $\theta_{NC}$  の値を求めるならば、 $D_C$  と  $D_N$  の値を求めることができ、これらの修正値を基に癌と癌以外の死亡率の解析を行うことができる。

この方法の定式化では(付録に詳述)、 $d_C$ ,  $d_N$ ,  $D_C$  および  $D_N$  の期待値に基づいて修正を行う

が、これらの期待値は  $\theta_{CN}$  および  $\theta_{NC}$  と共に、年齢、時間、被爆時年齢、都市、性、放射線量などの他の変数の関数である。したがって、考え方は本質的には同じであるが、解析に用いた誤分類の修正は上記のものより更に複雑である。この正式の方法を用いれば、尤度比テストを行うことができ、癌および癌以外の死亡率関数を扱うモデルのパラメータの最尤推定値を求めることができる。

DS86 コホート 本解析のコホートは、1950年10月1日から1985年12月31日までの期間中に追跡調査され、DS86線量<sup>26</sup>が算出されている寿命調査対象者86,520人から成る。このコホートには、最初のDS86コホート<sup>27</sup>より約10,000人多く含まれている。この増加は1989年に行われたDS86体系の拡大による。この拡大線量体系は線量推定Version 3と呼ばれ、藤田<sup>28</sup>によって報告されている。被爆都市(広島または長崎)、性、被爆時年齢(13区分)、グレイ単位の大腸の放射線吸収線量(10区分)、被爆後経過時間(7区分)により3,640個のセルに分類した人年および死亡数のデータを表に要約した。6 Gyを超える推定カーマ線量は6 Gyとした。検討する様々な種類の疾患および部位の臓器線量を代表するものとして、大腸線量を選択した。線量の中性子およびガンマ成分はRBE 1の定数で加重した。集団の62%は女性、67%は広島で被爆した者であった。表1および表2にこのコホートの他の特徴を要約する。

表1. 対象集団に関する情報の要約

線量(Gy)	対象者数	人年 (単位: 万)	平均線量 (Gy)	癌死亡数	癌以外の 死亡数
0.0	36,437	104.63	0.000	2600	10,201
0.01-0.05	27,054	78.00	0.0180	1949	7451
0.06-0.09	5305	15.27	0.0720	411	1509
0.10-0.19	6026	17.34	0.137	482	1701
0.20-0.49	6379	18.35	0.324	543	1785
0.50-0.99	3162	9.09	0.693	298	826
1.00-1.99	1552	4.53	1.35	170	369
2.00-2.99	371	1.04	2.35	40	86
3.00-3.99	149	0.406	3.52	19	51
≥4.0	85	0.237	4.43	10	23
合計/平均	86,520	248.9	0.113	6522	24,002



表2. 死亡年齢別のDS86コホートの(死亡診断書に基づく)死因の種類別死亡例数  
ならびに1961-75年の剖検例数

死亡年齢 (年)	癌死亡例数 1950-85年	剖検した癌の 死亡例数 1961-75年	癌以外の死亡例数 1950-85年	剖検した癌以外 の死亡例数 1961-75年
5-9	0	—	0	—
10-19	0	—	83	—
20-29	18	2	353	15
30-39	117	20	525	60
40-49	413	57	918	87
50-59	977	107	1801	144
60-69	1694	293	4045	472
70-79	2485	424	10,185	1127
80-89	796	81	5632	590
≥90	22	1	460	28
合計	6522	985	24,002	2523

放影研剖検データ 1961年1月から1975年12月までのABCC-放影研剖検プログラム<sup>29</sup>によって得たデータを用いて、誤分類の確率  $\theta_{CN}$  と  $\theta_{NC}$  のモデルを求めた。1961年以前の剖検データも存在するが、寿命調査集団の全死亡例の剖検を行うことを目的とした正式なプログラムは1961年まで開始されなかった。<sup>6</sup> 剖検プログラムは1980年代まで続いたが、剖検実施率は1975年以降急速に低下した。したがって、本報では1975年を最終年として扱う。DS86コホートの死亡例のうち剖検を行ったのは、1961年から1975年までは26%であるが、それ以前とそれ以降ではいずれも約6.5%であった。1961年に正式な剖検プログラムを開始してから行ったDS86コホートの剖検4,341例のうち3,703例、すなわち85%が1961年から1975年の間に実施されたものである。

剖検データを解析するために、死亡に至る一連の事象を引き起こした疾患または状態として定義される死亡診断書上の原死因を、剖検病理医が主要死因として指定した診断と比較した。<sup>29</sup> 死亡診断書上の死因の誤分類の大部分が、白血病を除く全癌と、事故を除く癌以外の疾患との間の誤りであることを剖検データは示している。例えば、剖検で癌と診断された1,173の死亡例のうち、死亡診断書では1例が白血病と書かれており、221例が癌以外と診断されていた。剖検で癌以外と診断された2,355の死亡例のうち5例は白血病、21例は事故、65例は癌となっていた。我々に関心があるのは主に癌以外と誤分類された癌症例の影響であり、白血病症例が癌以外として誤分類されることは他の癌を癌以外とするのに比べてまれであるので、白血病と事故による死亡は正しく分類されていると仮定し、専ら白血病を除く全癌

症例と事故を除く癌以外の症例間の誤分類に留意した。したがって、これ以降、“癌”という語は白血病を除く癌による全死亡を意味し、“癌以外”は癌および事故を除く原因による全死亡を意味する。より完璧なアプローチについては付録に説明するが、そのアプローチでは、全死因および死因間の誤分類をすべて検討する。しかし、死因をこれらの二つの群に大別することにより、問題をかなり単純化できる一方、観察された癌以外の線量反応が癌の誤分類によるか否かを的確に検討できる。DS86コホートで1961年から1975年までに実施された癌および癌以外の死亡例の剖検数を、年齢および死亡診断書の死因別に表2に示す。

どの死亡例を剖検するかという選択は、都市、性、時期および放射線量への強い依存性を示した。本解析では剖検例の選択が死亡診断書上の診断に依存しているという証拠は認められなかったが、最近のより総括的な誤分類の解析ではその影響があることが示されている。これについては以下に詳しく述べる。

癌および癌以外の誤分類率の推定値 剖検に基づいて診断された癌および癌以外の症例のうち、死亡診断書上で誤分類されていたものの割合を解析できたのは、DS86線量が算出されている者のみに限られている。これにはGMBOプログラム<sup>30</sup>を用いたロジスティック回帰法を使用した。死因をこの程度の細かさに分類した場合には、誤分類率は性、都市、剖検時期に有意には依存しなかった。更に、放射線量も誤分類率の有意な修飾因子ではなかった。しかしながら、標準誤差は大きかったが、癌の誤分類への線量の影響の点推定値は比較的大きかったことに留意すべきである。癌および癌以外の誤分類率は、ほぼ65歳を超える年齢では、死亡年齢に伴い統計的に有意な急速な変化を示した。いずれの誤分類についても、65歳未満の死亡年齢では年齢に対して不変であるが、65歳以降に死亡した者については年齢に依存する線形ロジスティックモデルは、データの本質的な特徴を表している。このモデルの式は次のとおりである。

$$\ln\left(\frac{\theta}{1-\theta}\right) = \gamma_0 + \gamma_1(\text{age} - 65)^+ \quad (2)$$

$(x)^+$ 量は $x < 0$ ではゼロ、 $x \geq 0$ では $x$ であり、 $\theta$ は $\theta_{NC}$ または $\theta_{CN}$ である。剖検データから推定したパラメータ値は、 $\theta_{CN}$ については $\gamma_0 = -2.19$ 、 $\gamma_1 = 0.100$ 、 $\theta_{NC}$ については、 $\gamma_0 = -3.25$ 、 $\gamma_1 = -0.040$ であった。死亡年齢全体で平均した場合には、これらから得た集団全体の粗誤分類率は癌では20%、癌以外では2.8%であった。主要解析ではこれらの誤分類モデルを用いた。それらを図1に示す。

感度解析には更に二つの誤分類モデルを用いた。まず、全体的な癌の誤分類率の過小評価に対する我々の結果の感度を検討するために、 $\gamma_0$ を $-1.374$ にまで増加させ、 $\gamma_1 = 0.100$ と

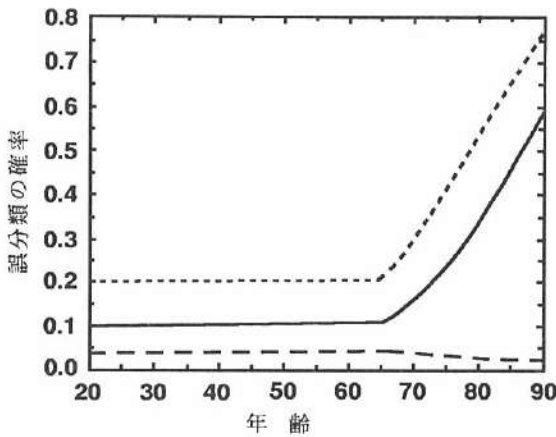


図1. 死亡年齢の関数としての誤分類の確率(等式2)。実線： $\gamma_0 = -2.19$ および $\gamma_1 = 0.100$ の $\theta_{CN}$ (癌全体の20%誤分類)，ダッシュ線： $\gamma_0 = -3.25$ および $\gamma_1 = -0.040$ の $\theta_{NC}$ (癌以外全体の2.8%誤分類)，点線： $\gamma_0 = -1.374$ および $\gamma_1 = 0.100$ の $\theta_{CN}$ (癌全体の36%誤分類)

した癌の誤分類関数を用いた解析を実施したところ、65歳以前に死亡した者の誤分類率はちょうど2倍になった。このモデルでは、癌の粗誤分類率は36%である。この関数も図1に示す。第2に、線量依存の $\theta_{CN}$ への影響を調べるために、(2)を拡大し、別の線量線形項および $\gamma_2$ を含む式を検討した。線形ロジスティック解析におけるこのパラメータの点推定値は0.23であり、95%信頼区間は(-0.15, 0.56)であった。感度解析のこの部分で用いた $\theta_{CN}$ のモデルでは、パラメータ値 $\gamma_0 = -2.19$ 、 $\gamma_1 = 0.100$ 、 $\gamma_2 = 0.23$ を用いた。表3はこの場合の癌の誤分類率と線量との関係を死亡年齢0-65歳および75歳について示す。

表3. 線量と死亡年齢の関数としての癌の誤分類率

年 齢	線 量(Gy)				
	0	0.5	1	1.5	2.0
0-65	10.1	11.2	12.4	13.7	15.1
75	23.4	25.5	27.8	30.1	32.6

以下に結果を示すのに好都合なように、癌の誤分類について用いた三つの異なる関数を各々、20%誤分類、36%誤分類、線量依存の20%誤分類と呼ぶ。

癌および癌以外の死亡率の回帰モデル 癌および癌以外の死亡率のモデル化にポアソン回帰法<sup>8</sup>を用いた。バックグラウンド率(0線量に被曝した人々の死亡率)を都市、性、被曝時年齢カテゴリー、および被曝後経過時間カテゴリーによって層化した。癌死亡率については、加法的相対リスクモデルを用いて、線量の影響をモデル化した。

$$RR_C(dose, city, sex, ager) = 1 + e^{a_1 city + a_2 sex + a_3(ager - 25)} \beta_1 dose \quad (3)$$

ただし, *sex* は女性の二値表示, *dose* は線量カテゴリー内のグレイ単位の平均腸線量, *age<sub>x</sub>* は被爆時年齢カテゴリー内の平均被爆時年齢である。予備解析では, *dose* の二乗の項を入れ, 誤分類について補正を行った結果として起こる癌線量反応の形状の変化を考慮できるようにした。いずれの解析でもこの二乗項は小さく統計的に有意でなく, それを除外しても関心の対象である結果には重要な影響は及ぼさないで, 線量に対して線形である癌相対リスクモデルの結果のみを示す。パラメータ  $\beta_1$  は広島で25歳のときに被爆した男性の1 Gyにおける過剰相対リスク(ERR, すなわち相対リスクから1を引いたもの)として解釈される。

癌以外の死亡率の加法的相対リスクモデルは次のとおりである。

$$RR_{NC}(dose) = 1 + \beta_1 dose + \beta_2 dose^2 \quad (4)$$

このモデルは線形-べきモデルと呼ぶ。癌以外の全死亡の予備解析では非線形が癌以外の線量反応の重要な特徴であることが示された。このような場合には従来, 線形-二次モデルを用いてきたが, 上記の線形-べきモデルを使うことにより低線量の発生率推定値がバックグラウンド発生率よりも小さくなる可能性がなくなる。後者の特徴は, 上昇曲率が非常に高いデータのモデル化に時折用いられる閾値モデルにも認められるが, そのようなモデルでは一定の閾値線量以下では放射線の影響はゼロと仮定する。線形-べきモデルを用いることにより, この閾値を推定する必要がなくなる。上記のべきモデルは癌以外の死亡率データによく当てはまる。この適合度は線形-二次モデルまたは線形-閾値モデルから得られる適合度についての統計的基準からみたのでは判断できない。パラメータ  $\beta_1$  は1 GyにおけるERRである。パラメータ  $\beta_2$  は曲率の程度を反映する。例えば,  $\beta_2 = 1$  は線量に対して純粋に線形のERRに対応し,  $\beta_2 = 2$  は純粋に二次の関係に対応する。

予備的解析でも, 癌以外の相対リスクへの性, 被爆時年齢, 時間の影響を示す証拠は認められなかった。上記の癌モデルと同様に, 都市, 性, 被爆時年齢に基づく線量反応の修正をして, 癌以外の死亡率の解析も行った。影響の修正を行ったモデルに基づく結果とモデル(4)に基づく結果の差異は無視できる程度であったので, モデル(4)に基づく結果のみを示す。

## 結 果

診断の誤分類が存在する場合の癌以外および癌の死亡率 表4左端の縦2列に, 誤分類について補正を行った場合と行わない場合の, 癌以外および癌の死亡率解析を要約する。表の最初の横3列に癌以外の線量反応の次の三つの異なる仮説検定の尤度比  $\chi^2$  統計量を示す。

表4. 誤分類について補正をした場合としなかった場合の癌以外の線量反応の尤度比テストおよび癌以外(等式4)と癌(等式3)の相対リスク関数のパラメータ推定値

癌以外の線量 反応テスト	誤分類の 補正なし	20%誤分類の 補正	感度解析	
			36%誤分類の 補正	$\theta_{CN}$ に線量影響がある 20%誤分類の補正
Null vs. L-P $x_2^2$ (p値)	18.7 (<.001)	11.3 (<.004)	5.9 (.052)	4.3 (.12)
Null vs. L $x_1^2$ (p値)	15.6 (<.001)	8.7 (.003)	3.8 (.051)	2.9 (.089)
L vs. L-P $x_1^2$ (p値)	3.1 (.078)	2.6 (>.10)	2.1 (>.10)	1.4 (>.10)
$\beta_1$	0.0630 (0.0317)	0.0498 (0.0320)	0.0336 (0.0326)	0.0308 (0.0315)
$\beta_2$	1.560 (0.444)	1.631 (0.554)	1.793 (0.803)	1.678 (0.863)
ERRの%変化(対補正なし)	-	-21	-47	-51
癌以外の過剰死亡数(%変化)	129	99 (-23)	64 (-50)	62 (-52)
$\alpha_1$	-0.676 (0.350)	-0.712 (0.327)	-0.674 (0.305)	-0.650 (0.289)
$\alpha_2$	0.710 (0.294)	0.668 (0.267)	0.602 (0.248)	0.548 (0.231)
$\alpha_3$	-0.0416 (0.00946)	-0.0441 (0.00870)	-0.0445 (0.00810)	-0.0395 (0.00773)
$\beta_1$	0.494 (0.129)	0.553 (0.131)	0.556 (0.121)	0.675 (0.137)
ERRの%変化(対補正なし)	-	12	13	37
癌の過剰死亡数(%変化)	274	317 (16)	369 (35)	374 (36)

注. ERR = 過剰相対リスク, L = 線形モデル, およびL-P = 線形-べきモデル

1) Nullモデル(すなわち線量依存なし)対 線形-べきモデル, 2) Nullモデル 対 線形モデル, 3)線形 対 線形-べきモデル。表4の残りの横の列には, 癌については線量に対して線形のモデル, 癌以外に対しては線形-べきモデルに基づく, 癌および癌以外の相対リスク関数のパラメータ推定値を示す。過剰相対リスクの変化の割合, ならびにDS86コホートにおける放射線被曝による過剰症例の推定数およびその変化の割合も示す。

誤分類について補正をしない場合, 癌以外の死亡率の線形-べき線量反応は有意で,  $\chi^2_2 = 18.7$  ( $p < .001$ )である。この  $\chi^2$  は相対リスクの線形増加の検定  $\chi^2_1 = 15.6$  ( $p < .001$ )に大きく起因し得るが, 非線形性を考慮に入れることにより幾分かの改善ができる。20%誤分類について補正を行っても, 線形-べき線量反応は依然として有意であり( $p < .004$ ), この場合も相対リスクの線形増加の検定に起因するところが大きい。しかし, 癌以外の線量反応における非線形性の程度に影響を与えるパラメータ  $\beta_2$  の推定値は, いずれの場合も1.5より大きく, このことはある意味ではデータに適合するモデルは純粋な線形モデルよりも純粋な二次モデルに近いことを示していることに留意すべきである。誤分類について補正することにより, 1 Gyにおける癌以外のERRの推定値は0.0630から0.0498に, すなわち21%減少する。癌以外の過剰死亡数は, 129から99に, すなわち23%減少する。癌のERR(25歳で広島で被曝した男性について推定)は0.494から0.553に増加(12%)し, 癌の過剰死亡数は274から317に増加(16%)する。この場合, 誤分類の補正により, 正味832例が癌以外から癌に再分類され, 癌以外の死亡数は3.5%減少し, 癌の死亡数は12.8%増加する。

感度解析 線量に依存する36%誤分類と20%誤分類の補正を行った感度解析の結果を表4右端の縦2列に示す。初めの感度解析では, 36%誤分類について補正すると, 癌以外の死亡率の線量反応は以前より平坦になり, 有意性は境界線上まで減少した( $p = .052$ )。この場合, 癌以外の推定ERRは47%, 癌以外の過剰死亡数は50%減少する。癌の推定ERRは13%増加するが, 癌の過剰死亡数は35%増加する。癌の誤分類率を2倍にしたところ, 癌のERRはわずかしこ増加しなかったが, 誤分類補正によって癌の推定バックグラウンド率が増加するので, 癌の過剰死亡数はかなり増加した。第2の感度解析では, 線量に依存する20%誤分類について補正すると, 癌以外の線量反応の有意性は認められなくなった( $p = .12$ )。補正をしない場合に比べて, 癌以外のERRは52%減少し, 癌のERRは37%増加する。これに対応する過剰死亡数の変化は, 癌以外が52%の減少, 癌が36%の増加である。

## 考 察

寿命調査集団の癌以外および癌の死亡率の解析は, 癌以外の死亡率について認められた有意な線量反応を, 癌を癌以外とする線量に依存しない誤分類によって簡単には説明できない



ことを示す。放影研の剖検プログラムから推定される誤分類の割合、すなわち癌の20%誤分類と癌以外の2.8%誤分類について補正すると、癌以外の線量反応の有意水準が低下し、推定過剰相対リスクは18%、癌以外の過剰死亡数は23%減少する。しかし、癌以外の死亡率線量反応は依然として有意である( $p = .004$ )。また、誤分類について補正すると癌の推定ERRは12%増加し、過剰癌死亡数は16%増加する。

癌以外の線量反応の統計的有意性は全体的な癌の誤分類率の過小評価に対しては、比較的感受性が低いことが感度解析により明らかになった。全体的な癌の誤分類率を36%に倍増しても、癌以外の線量反応は依然として有意性の境界線上にあった( $p = .052$ )。しかし、これらの解析は、癌の誤分類が線量と正の関係をもち得るか否かに対しては、癌以外の線量反応の有意性が高い感度をもつことも示した。1 Gyで癌の誤分類率の約20%の増加に相当する線量への依存を仮定した場合、癌以外の線量反応は、もはや有意でなかった。しかし、その場合、1Gyでの癌の推定ERRは37%増加した。

この最後の感度解析では、癌症例の誤分類は、低線量群より高線量群で系統的により頻繁に起きると仮定していることに留意すべきである。この仮定が成立するためには、死因を診断する医師が、個人の被曝レベルについて何らかの知識があり、その情報により意識的にまたは無意識的に診断基準を変え、高線量の症例でより頻繁に癌症例が癌以外と誤診されることが必要と考えられる。死亡診断を行う人々は、線量と相関関係をもつ爆心地からの距離について大まかに知っているかもしれないが、個人の線量について具体的な情報を知っているとは考えられない。高線量の放射線に被曝した症例では癌以外よりもむしろ癌に起因する死亡が多いと考えられるので、この種の誤分類の線量への依存は直感に反すると思われる。放射線誘発癌は診断が最も困難な癌でもあるので、癌の種類を考慮に入れないければ、誤分類の線量反応が観察される可能性もある。現在行っている誤分類のより詳細な解析により、誤分類の確率の線量反応をこれまで述べてきたような解析に含めるべきか否かが明確になるかもしれない。

癌および癌以外の誤分類率が、都市、性、または被曝後経過時間に有意に依存しないことが剖検データから判明した。しかし、誤分類率は死亡年齢に極めて強く依存しており、このことは我々のモデルで考慮されている。剖検データから得られる線量への誤分類率の依存に関する証拠は弱い。どの患者が剖検に選ばれたかが都市、性、被曝後経過時間、線量、被曝時年齢に強く依存することを示す証拠はあったが、これらの影響を考慮に入れた後に死亡診断書上の死因に依存することを示す証拠はなかった。ただし、最近の誤分類に関する解析はそのような影響を示唆している。剖検率が最初の診断に大きく依存するならば、癌の誤分類率の我々の推定値に偏りが生じ得るので、この点は重大な問題である。しかし我々は、感度解

析によって、癌の誤分類率を過小評価することによる影響について検討した。

我々の剖検データの解析は総括的なものではなく、本報の特定の目的に合わせて実施したものであることに留意しなければならない。最近のより総括的な誤分類の解析では、より正確な誤分類の補正が可能な領域があることが指摘されている。例えば、癌の推定誤分類率は剖検場所(ABCC-放影研または病院)と死亡場所(病院または家)に依存すると思われる。前者について言えば、死亡診断書を完成させる際に、病院の剖検結果を考慮に入れたかもしれない。ABCC-放影研の剖検記録のみに基づく癌の全体的な誤分類率の推定値は、我々が用いた20%よりもむしろ25%に近かった。これは依然として我々が感度解析で考えた値の範囲内に十分入っている。後者について言えば、病院での死亡は家での死亡に比べて、より正確に診断される。剖検対象者の家と病院での死亡の分布は一般集団と異なっているかもしれない、そのことは我々が用いている“平均”誤分類率が適切でないことを示唆しているので、将来の解析においては、この区別をすることが重要であるかもしれない。

本報では、推定誤分類関数を固定したものとして扱い、癌と癌以外の死亡率を推定したが、死亡率データと剖検データを総合したデータのすべてを使った誤分類関数の推定はしなかった。総合データを用いて死亡率関数と誤分類率を合わせて推定することも可能であるが、本報では癌と癌以外の死亡率線量反応に関する推定への誤分類の影響に焦点を当てた。そのために、剖検データを用いることによって、寿命調査における誤分類のメカニズムをみるための妥当と思われるモデルを推定し、これらのモデルが固定していると考え、誤分類に関する補正の基盤として用いた。

特定の死因(心疾患など)の誤分類に関する補正をより正確にするために、単に二つのカテゴリー、すなわち心疾患カテゴリーと他の全ての死因のカテゴリーを作り、これら二つのカテゴリー間の誤分類の確率を推定し、上記の方法を当てはめることもできる。この方法の利点は、当てはめるのが比較的簡単なことである。欠点は、非常に様々な死因を含む広い包括的なカテゴリーは生物学的に解釈できないかもしれないことと、関心のある各死亡カテゴリーについて別個の解析を行わなければならないことである。その代わりに、付録に示すように、類似する死因をひとまとめにし、二つ以上の区分にこの手順を拡大することもできる。この実行はより複雑であるが、解釈の可能性および内部での一貫性の点で利点があるかもしれない。癌以外の特定の死因による放射線量との関連が、癌の誤分類によって説明できるかを究明する上で、また死亡診断書上の診断の誤分類に特に影響を受けやすい肝臓癌など特定の種類の癌について、放射線リスクの推定値をより正確にする上で、いずれの方法も有益であろう。

## 参考文献

1. Shimizu Y, Kato H, Schull WJ, Hoel DG: Life Span Study Report 11. Part 3. Noncancer mortality, 1950–1985, based on the revised doses (DS86). RERF TR 2-91
2. Beebe GW, Ishida M: Joint JNIIH–ABCC study of life span of atomic bomb survivors. Research plan. ABCC TR 4-59
3. Shimizu Y, Kato H, Schull WJ: Studies of the mortality of A-bomb survivors. 9. Mortality, 1950–1985: Part 2. Cancer mortality based on the recently revised doses (DS86). Radiat Res 121:120–41, 1990
4. National Research Council: *Health Effects of Exposure to Low Levels of Ionizing Radiation* (BEIR V). Washington, D.C., National Academy Press, 1990
5. Hollingsworth JW, Beebe GW: ABCC–JNIIH Adult Health Study: Provisional research plan. ABCC TR 9-60
6. Zeldis LJ, Matsumoto YS: ABCC–JNIIH pathology studies, Hiroshima–Nagasaki, provisional research plan. 1. Description and scope of program. ABCC TR 4-61
7. Dempster AP, Laird NM, Rubin DB: Maximum likelihood from incomplete data via the EM algorithm. J Royal Stat Soc Ser B 39:1–38, 1977
8. Clayton D: The analysis of event history data: A review of progress and outstanding problems. Stat Med 7:819–41, 1988
9. Dinse GE: Nonparametric prevalence and mortality estimators for animal experiments with incomplete cause-of-death data. J Am Stat Assoc 81(394):328–36, 1986
10. Kodell RL, Chen JJ: Handling cause of death in equivocal cases using the EM algorithm. Commun Stat Theory Methods 16(9):2565–85, 1987
11. Dinse GE: Estimating tumor incidence rates in animal carcinogenicity experiments. Biometrics 44:405–15, 1988
12. Archer LE, Ryan LM: Accounting for misclassification in the cause-of-death test for carcinogenicity. J Am Stat Assoc 84(407):787–91, 1989
13. Espeland MA, Platt OS, Gallagher D: Joint estimation of incidence and diagnostic error rates from irregular longitudinal data. J Am Stat Assoc 84(408):972–9, 1989
14. Duffy SW, Rohan TE, Day NE: Misclassification in more than one factor in a case-control study: A combination of Mantel-Haenszel and maximum likelihood approaches. Stat Med 8(12):1529–36, 1989
15. Elton RA, Duffy SW: Correcting for the effect of misclassification bias in a case-control study using data from two different questionnaires. Biometrics 39(3):659–63, 1983
16. Flegal KM, Brownie C, Haas JD: The effects of exposure misclassification on estimates of relative risk. Am J Epidemiol 123(4):736–51, 1986
17. Greenland S: Statistical uncertainty due to misclassification: Implications for validation substudies. J Clin Epidemiol 41(12):1167–74, 1988
18. Greenland S: On correcting for misclassification in twin studies and other matched-pair studies. Stat Med 8(7):825–9, 1989
19. Greenland S, Robins JM: Confounding and misclassification. Am J Epidemiol 122(3):495–506, 1985

20. Hsieh CC, Walter SD: The effect of non-differential exposure misclassification on estimates of the attributable and prevented fraction. *Stat Med* 7(10):1073–85, 1988
21. Savitz DA, Baron AE: Estimating and correcting for confounder misclassification. *Am J Epidemiol* 129(5):1062–71, 1989
22. Green MS: Use of predictive value to adjust relative risk estimates biased by misclassification of outcome status. *Am J Epidemiol* 117(1):98–105, 1983
23. White E: The effect of misclassification of disease status in follow-up studies: implications for selecting disease classification criteria. *Am J Epidemiol* 124(5):816–25, 1986
24. Chen TT: A review of methods for misclassified categorical data in epidemiology. *Stat Med* 8:1095–106, 1989
25. Walter SD, Irwig LM: Estimation of test error rates, disease prevalence and relative risk from misclassified data: A review. *J Clin Epidemiol* 41(9):923–37, 1988
26. Roesch WC (ed.): *US—Japan Joint Reassessment of Atomic Bomb Radiation Dosimetry in Hiroshima and Nagasaki: Final Report* (Volumes 1 and 2). Hiroshima, Radiation Effects Research Foundation, 1987
27. Preston DL, Pierce DA: The effects of changes in the dosimetry on cancer mortality risk estimates in the atomic bomb survivors. *Radiat Res* 114:437–66, 1988
28. Fujita S: Versions of DS86. RERF Update 1(2): 3, 1989 (Hiroshima, Radiation Effects Research Foundation)
29. Yamamoto T, Moriyama IM, Asano M, Guralnick L: RERF pathology studies, Hiroshima and Nagasaki. Report 4: The autopsy program and Life Span Study, January 1961–December 1975. RERF TR 18-78
30. Preston DL, Lubin JH, Pierce DA: *Epicure: User's Guide*. Seattle, HiroSoft International, 1990

## 付 録

## ポアソン回帰における癌および癌以外の誤分類に関する補正

死亡率関数  $\lambda(t)$  は共変量の時間変動ベクトル  $Z(t)$  に依存すると仮定し、 $\lambda_k$  値をとる区分的に一定の関数によってこの死亡率関数を近似する。ただし、 $k = 1, \dots, K$  は  $Z(t)$  空間の分割の  $K$  個のセルの一つを示す。次に、データの近似尤度は

$$\prod_{k=1}^K \lambda_k^{D_k} \exp[-P_k \lambda_k]$$

に比例する。ただし、 $P_k$  は  $k$  番目のセルのすべての人々の人年の合計であり、 $D_k$  はそのセルの人々に見られた死亡数である。<sup>8</sup> 実際は、階段関数  $\lambda_k$  は連続関数  $\lambda(z_k)$  によって表される。ただし、共変量ベクトル  $Z(t)$  は  $z_k$ 、 $k = 1, \dots, K$  の  $K$  個の値の一つで打ち切る。 $\lambda(z_k)$  の様々な非線形モデルは、特にこの問題用に設計されたコンピュータプログラム AMFIT<sup>30</sup> を用いると簡単に当てはめることができる。

次に、上記の分割されたセルの一つを考え、下記の量を定義する。(表示を簡素化するために、添字  $k$  および独立変数を省略した。)

- $P$  人年数
- $D_C$  真の癌死亡数
- $D_N$  真の癌以外の死亡数
- $d_C$  観察癌死亡数(死亡診断書による)
- $d_N$  癌以外の観察死亡数(死亡診断書による)
- $D$  合計死亡数 =  $D_C + D_N = d_C + d_N$
- $X_C$  誤分類した真の癌死亡数
- $X_N$  誤分類した真の癌以外の死亡数
- $\lambda_C$  癌死亡率
- $\lambda_N$  癌以外の死亡率
- $\theta_{CN}$  癌の誤分類率-死亡診断書で真の癌症例が癌以外に誤分類される確率
- $\theta_{NC}$  癌以外の誤分類率-死亡診断書で真の癌以外の症例が癌と誤分類される確率

$\theta_{CN}$  と  $\theta_{NC}$  は各々癌の診断に関する偽陰性率と偽陽性率であることに留意する。また、疫学で通常用いる用語では、 $(1 - \theta_{CN})$  と  $(1 - \theta_{NC})$  が各々癌と癌以外の発見率である。

	真の癌 症例	真の癌以外の 症例	
死亡診断書上の 癌の診断	$\frac{D_C - X_C}{\lambda_C + \lambda_N} (1 - \theta_{CN})$	$\frac{X_N}{\lambda_C + \lambda_N} \theta_{NC}$	$d_C = D_C - (X_C - X_N)$
死亡診断書上の 癌以外の診断	$\frac{X_C}{\lambda_C + \lambda_N} \theta_{CN}$	$\frac{D_N - X_N}{\lambda_C + \lambda_N} (1 - \theta_{NC})$	$d_N = D_N + (X_C - X_N)$
	$D_C$	$D_N$	$D$

図 A1. 各分類カテゴリーの死亡例数 D および多項確率。欄外は合計死亡数

図 A1 は真の原死因と死亡診断書上の診断によって死亡を交差分類した場合の各セルの死亡数および多項確率を示す。欄外に合計死亡数も示す。

死亡診断書上の癌の診断が実際は癌以外である確率 ( $\theta'_{CN}$ ) および死亡診断書上の癌以外の診断が実際は癌である確率 ( $\theta'_{NC}$ ) は以下のように簡単に示すことができる。

$$\theta'_{CN} = \frac{\lambda_N \theta_{NC}}{\lambda_C (1 - \theta_{CN}) + \lambda_N \theta_{NC}}$$

$$\theta'_{NC} = \frac{\lambda_C \theta_{CN}}{\lambda_C \theta_{CN} + \lambda_N (1 - \theta_{NC})} \quad (A.1)$$

一般的にこれらの確率は  $\lambda_C$  と  $\lambda_N$ ,  $\theta_{CN}$  と  $\theta_{NC}$  に影響を与える線量とその他の変数に依存する。 $(1 - \theta'_{CN})$  および  $(1 - \theta'_{NC})$  量は各々、癌および癌以外の確認率と呼ばれることがある。

真の(観察されない)死亡頻度  $D_C$  と  $D_N$  および誤分類した症例  $X_C$  と  $X_N$  に関して、 $K$  次元の人年表の一つのセルの完全なデータ尤度は

$$L(D_C, D_N, X_C, X_N | \lambda_C, \lambda_N, \theta_{CN}, \theta_{NC}) = \left[ (P\lambda_C)^{D_C} (P\lambda_N)^{D_N} e^{-P(\lambda_C + \lambda_N)} \right] \cdot \left[ \theta_{CN}^{X_C} (1 - \theta_{CN})^{D_C - X_C} \theta_{NC}^{X_N} (1 - \theta_{NC})^{D_N - X_N} \right] \quad (A.2)$$

に比例する。(A.2)の右側の最初の項は二つの独立ポアソン確率変数の結合尤度に比例し、第2の項は二つの独立二項確率変数の尤度に比例することに留意する。癌および癌以外の死亡率の二つのポアソン尤度、ならびに各種の誤分類それぞれの二つの二項尤度を別個に最大にすることによって、この尤度を最大にする。



$X_C$  および  $X_N$  が、したがって  $D_C$  および  $D_N$  も、観察されない場合、観察されるデータの尤度は

$$L(d_C, d_N | \lambda_C, \lambda_N, \theta_{CN}, \theta_{NC}) = \{P[\lambda_C(1-\theta_{CN}) + \lambda_N\theta_{NC}]\}^{d_C} \{P[\lambda_C\theta_{CN} + \lambda_N(1-\theta_{NC})]\}^{d_N} e^{-P(\lambda_C + \lambda_N)} \quad (A.3)$$

に比例する。ただし、角括弧内の量は観察された癌および癌以外の死亡のポアソン率であり、これは誤分類の確率と合成した真の癌および癌以外の死亡のポアソン率である。全体の完全データ尤度と観察可能なデータ尤度は、 $K$  個のセルの各々について、それぞれ (A.2) および (A.3) のような項の積である。

EM アルゴリズム<sup>7</sup> を用いて、(A.3) に基づく対数尤度を最大にすることができる。E では、(A.2) の  $D_C, D_N, X_C$  および  $X_N$  は観察データ  $d_C$  および  $d_N$  から得た期待値に置き換える。図 A1 からこれらの期待値は次のようになることがわかる。

$$\begin{aligned} E(D_C | d_C, d_N) &= d_C(1-\theta'_{CN}) + d_N\theta'_{NC} \\ E(D_N | d_C, d_N) &= d_C\theta'_{CN} + d_N(1-\theta'_{NC}) \\ E(X_C | d_C, d_N) &= d_N\theta'_{NC} \\ E(X_N | d_C, d_N) &= d_C\theta'_{CN} \end{aligned} \quad (A.4)$$

M ステップでは、(A.4) から求めた期待値を用いて尤度 (A.2) を最大にする。(A.3) に基づき尤度を算出することにより尤度比検定を行うが、これは簡単に実施できる。

考察で言及したように、この最大化の方法では、原則として、副次群について剖検データを直接使うことができる。剖検情報が存在するデータ部分については、(A.2) 式の  $D_C, D_N, X_C, X_N$  は判明している。したがって、EM アルゴリズムの E ステップは不必要である。剖検情報がない死亡例については、上記のとおり EM アルゴリズムを当てはめる。しかし本解析では、剖検データから  $\theta_{CN}$  および  $\theta_{NC}$  を推定するが、それらを既知の関数として扱う。この場合、最大化には等式 (A.2) のポアソンに類似する項のみが必要である。

#### 複数の死因への拡大

概念的には、複数の死因があり、それらの中に誤分類がある場合に、この手順を拡大してそれに適用できるが、実際に行うことは考えるほど容易ではない。死因が  $M$  個あると考え、 $\lambda_i$  を死因  $i, i=1, \dots, M$  の死亡率関数とする。更に、真の死因は  $i$  であるが死亡診断書上では死因が  $j$  と分類されている人々の割合を  $\theta_{ij} (i, j=1, \dots, M)$  とする。同様に、 $D_i$  と  $d_i$  を死因  $i$  の真の

死亡数と観察死亡数とし、 $X_{ij}$ を死因 $j$ に分類されたが実は $i$ である死亡例数とする。この場合、上記(A.2)に対応する尤度は $X_i = D_i$ および $\theta_i = 1$ という条件で

$$L(D_i, X_{ij}, i, j = 1, \dots, M) = \prod_{i=1}^M (P\lambda_i)^{D_i} e^{-P\lambda_i} \prod_{i,j=1}^M \theta_{ij}^{X_{ij}} \quad (\text{A.5})$$

に比例する。ただし、添字“ $\cdot$ ”は添数の総和を表す。(A.3)に対応する尤度は

$$L(d_i, i = 1, \dots, M) = \prod_{i=1}^M \left[ \sum_{j=1}^M \lambda_j \theta_{ji} \right]^{d_i} e^{-P\lambda_i} \quad (\text{A.6})$$

に比例する。ただし、 $d_i = X_i$ である。等式(A.1)に対応する量は

$$\theta'_{ij} = \frac{\lambda_j \theta_{ji}}{\sum_{k=1}^M \lambda_k \theta_{ki}} \quad (\text{A.7})$$

であり、等式(A.4)に対応するのは

$$\begin{aligned} E(D_i | d_1, \dots, d_M) &= \sum_{j=1}^M d_j \theta'_{ji} \quad \text{および} \\ E(X_{ij} | d_1, \dots, d_M) &= d_j \theta'_{ji} \end{aligned} \quad (\text{A.8})$$

である。

これらの等式を用いることにより、EMアルゴリズムを前と同様に続けることができる。