

**A METHOD FOR ANALYSIS OF CANCER INCIDENCE IN
ATOMIC BOMB SURVIVORS, WITH APPLICATION TO ACUTE LEUKEMIA**

原爆被爆者における癌発生率の解析法—その急性白血病への応用

DONALD A. PIERCE, Ph.D.

DALE L. PRESTON, Ph.D.

TORANOSUKE ISHIMARU, M.D., M.P.H. 石丸寅之助



RADIATION EFFECTS RESEARCH FOUNDATION
財団法人 放射線影響研究所

A Cooperative Japan – United States Research Organization

日米共同研究機関

In the continued interest of accurately defining the late effects of the atomic bombs, the qualitative and quantitative characteristics of the A-bomb radiation exposure doses are periodically refined. If warranted by future dose assessments, the data reported here will be reanalyzed and subsequently reported.

原爆の後影響を引き続いて正確に究明する目的をもって、原爆放射線被曝線量の質的・量的特質について定期的に改良を加えている。今後線量評価によって、その必要性が起これば、本報の資料を再解析の上、改めて報告する。

RERF TECHNICAL REPORT SERIES

放影研業績報告書集

The RERF Technical Reports provide the official bilingual statements required to meet the needs of Japanese and American staff members, consultants, and advisory groups. The Technical Report Series is not intended to supplant regular journal publication.

放影研業績報告書は、日米専門職員、顧問、諮問機関の要求に応えるための日英両語による公式報告記録である。業績報告書は通例の誌上発表論文に代わるものではない。

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.

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もので、その経費は日米両政府の平等分担により、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運営されている。



RADIATION EFFECTS RESEARCH FOUNDATION
財団法人放射線影響研究所

A METHOD FOR ANALYSIS OF CANCER INCIDENCE IN ATOMIC BOMB SURVIVORS, WITH APPLICATION TO ACUTE LEUKEMIA

原爆被爆者における癌発生率の解析法—その急性白血病への応用

DONALD A. PIERCE, Ph.D.; DALE L. PRESTON, Ph.D.;
TORANOSUKE ISHIMARU, M.D., M.P.H. (石丸寅之助)

Department of Epidemiology & Statistics
疫学統計部

SUMMARY

Some statistical methodology developed for the analysis of the effect of radiation on cancer incidence is described. The methods presented make use of grouped data for estimation of and testing hypotheses about temporal aspects of the radiogenic effects on cancer incidence. The suggested approach is reviewed by its application to recent data on acute leukemia incidence in the RERF Life Span Study sample. The results of these analyses suggest that, although the temporal pattern of the radiogenic excess differs markedly by age at the time of the bomb (ATB), there is no statistically significant evidence of differences in total excess leukemia by age ATB throughout the study period.

INTRODUCTION

The purpose here is to describe some statistical methodology developed for studying the effect of radiation exposure on cancer incidence in the Japanese atomic bomb survivors. Various reports, over the years, have discussed not only the general dose response, but apparent interactions between factors such as age ATB, sex, city, dose, and time since exposure. It is important to provide methods for significance tests and systematic estimation of these relationships.

要約

放射線が癌発生率に及ぼす効果を解析するために開発された統計学的方法について述べた。この方法では、区分けしたデータを用いて、癌発生率に対する経時的な放射線誘発効果を推定し、その仮説を検討した。放影研の寿命調査対象集団における急性白血病の発生率に関する最近のデータにこの方法を適用することにより、その検討を行った。解析結果の示唆するところによれば、放射線誘発による過剰リスクの経時的パターンは、被爆時年齢によって顕著な差異が見られるが、調査期間全体の白血病の総過剰は被爆時年齢によって統計的有意差はない。

緒言

本論文の目的は、放射線が日本人原爆被爆者の癌発生率に及ぼす効果を研究するために開発された統計学的方法について述べることである。長年にわたり、数多くの論文で一般的な線量反応のみならず、原爆時年齢、性、都市、線量及び被曝後経過年数のような因子間の明らかな相互作用について論じられてきた。これらの関係の有意性検定及び系統的な推定の方法を提供することが重要である。

The emphasis in the present study is on investigation of the temporal nature of excess cancer risks, an aspect which has received considerable attention, but relatively little in the way of formal statistical analysis. The problem can be classed in the survival analysis category; however, the perspective is perhaps somewhat different than usual since the sample size is very large and the events of interest are rare. Also, censoring prior to the end of the follow-up period is essentially all due to bona fide competing risks, that is, to other causes of death.

Most analyses of these data are done by specific cancer type, since both the radiation effects and the background rates differ by site. Acute leukemia incidence was chosen for analysis in this report, but the methods are currently being used for other types of cancer as well. Although the example is central to the paper, the primary intent here is to describe the methodology, rather than to report on actual leukemia risks. Additional details related to dosimetry, case ascertainment, and other matters would need more attention in a proper report on leukemia incidence. Although there are a number of programs at RERF,¹ one of central importance is the Life Span Study (LSS), which involves the lifetime follow-up of 108,670 residents of Hiroshima and Nagasaki. This study provides the data base for the example here.

The LSS extended cohort consists of most survivors who were exposed within 2,500 m from the hypocenter and who were still living in the cities in 1950, group-matched by age and sex with a control sample chosen from those who were exposed between 2,500 and 10,000 m from the hypocenter. The sample also includes a group of people who were not in the city ATB. Further details of the selection of the sample are given elsewhere.^{2,3}

Persons in the exposed group were interviewed shortly after 1950 to determine information on location and shielding ATB. From this information radiation exposure level estimates (herein called doses) would eventually be computed. The dose estimates used here were computed using the 1965 air dose,⁴ and are referred to as the T65DR kerma doses.⁵ Major efforts to reevaluate these estimates are currently underway (see for example, US Department of Energy, 1982⁶). It is too early to predict even in which

この研究の重点は、かなり注目されてきた点ではあるが、正式な統計学的解析では注目されることが比較的少ない癌の過剰リスクの経時的特性を調べることにある。この問題は生存率解析の範ちゅうに分類されるが、集団が極めて大きく関心がもたれる事象はまれであるので、恐らく通常とは幾分異なる見通しとなるであろう。また、追跡調査期間終了前に行われる調査は、本質的にはすべて真の競合リスク、つまり他の死因のためである。

放射線効果及びバックグラウンド比率はいずれも部位により異なるので、これらのデータの解析はほとんど癌の型ごとに行う。本報の解析には急性白血病の発生率を選んだが、この方法は現在他の型の癌にも用いられている。急性白血病が本論文の中心であるが、主な意図は実際の白血病リスクを報告することではなくて、方法論を述べることである。線量測定、症例確認その他に関する詳細については、白血病の発生率に関する適当な論文で更に注意深く検討する必要がある。放影研では数多くの研究¹が進められているが、最も重要なものの一つは、広島及び長崎の住民108,670人を生涯にわたって追跡調査する寿命調査(LSS)である。本報に取り上げた例のデータベースはこの調査から提供された。

寿命調査拡大コホートは、爆心地から2,500 m以内で被爆し、1950年に市内に居住していたほとんどの被爆者、及びそれと年齢、性を一致させた爆心地から2,500~10,000 mの間にいた者から抽出された対照集団から成る。この集団は原爆時市内不在者群も含む。集団の抽出については別に詳しく述べられている。^{2,3}

被爆群については1950年以後間もなく、原爆時の位置及び遮蔽状況に関する面接調査を行った。この資料から、最終的には放射線被曝レベルの推定値(本報では線量と呼ぶ)が計算されることになった。本報で用いた線量推定値は1965年空気線量を用いて計算されたもので、⁴ T65DR kerma 線量と呼称されている。⁵ 現在、これらの推定値を再評価する努力がなされている(例えば、米国エネルギー省 1982年版⁶を参照)。線量反応の基本勾配の変化の方向を予想する

direction the basic slope of the dose response will change, but certainly the parameter estimates given here must be taken with some skepticism in regard to the actual levels of risk. However, the inferences discussed here about the temporal patterns of risk, and their interaction with factors such as age ATB, are likely to stand up under dose revisions. As the primary goal here is to develop statistical methodology suitable for these data, it seems advisable to proceed despite the uncertainty of the dosimetry.

Complete ascertainment of mortality and cause of death for the LSS extended sample is made possible by a traditional Japanese national system of family records (koseki). Further information on cancer incidence, as opposed to mortality, has been obtained from tumor registries and in cooperation with those engaged in medical care. The inferences presented here must necessarily pertain to cancer incidence after 1950, the time when the LSS sample was selected. It seems clear that there was substantial radiation-induced cancer incidence, especially leukemia, prior to that time.

A great deal of rather detailed data on cancer mortality in the LSS sample has been published in RERF Technical Reports at 4-year intervals. Data through 1978 appear in LSS Report 9, and its supplementary tables.^{5,7} These data should be a useful resource for statisticians, in both research and teaching. A great deal of useful background information, and results from analyses of LSS data, can be found in the report of the Advisory Committee on the Biological Effects of Ionizing Radiation (BEIR-3)⁸ and in the National Council on Radiation Protection and Measurements Report 64 (NCRP-64)⁹.

Although the example here pertains to incidence rather than mortality, similar calculations can be carried out on the mortality data of LSS Report 9. The format of the data used for the calculations here proved too voluminous to include in this paper, but a tabulation with broader age ATB categories, which yields very similar results, is given in the Appendix.

Much of the formal statistical analysis of these data has consisted of relating the cancer rates over the entire follow-up period to dose level, regardless of the actual dates of death or diagnosis. In fact, this turns out to be a

ことさえ時期尚早であるが、本報で示すパラメーター推定値については、実際のリスクレベルの点において明らかに幾分問題を含むと考えられる。しかしながら、本報で論ずるリスクの経時的パターンに関する推論、及びそのパターンと原爆時年齢などの因子間の相互作用は、線量の改訂後も有効であろう。本報での主目標は、これらのデータに適する統計学的方法を開発することであるので、線量測定が不確実であっても研究を進めるのが望ましいと思われる。

日本の伝統的な戸籍制度により、寿命調査拡大群の死亡率及び死因が完全に確認できる。死亡率に対して、癌発生率の更に詳しい資料は腫瘍登録及び医療の従事者の協力によって得られる。本報で述べる推論は、必然的に寿命調査対象群が抽出された1950年以後の癌発生率に関するものである。それ以前にも、放射線誘発性の癌、特に白血病の発生率が高かったということは明らかであると思われる。

寿命調査対象群の癌死亡率に関する数多くの、かなり詳細なデータが4年ごとに放影研業績報告書として発表されている。1978年までのデータは寿命調査第9報及びその付録表^{5,7}に示されている。これらのデータは、統計学者にとって研究及び教材に有益な資料であるに違いない。多量の有用な背景資料及び寿命調査データの解析結果は、電離放射線の生物学的影響に関する諮問委員会の報告書 (BEIR-3)⁸ 及び米国放射線防護委員会報告書64 (NCRP-64)⁹ に掲載されている。

ここに示した例は死亡よりもむしろ発生に関するものであるが、寿命調査第9報の死亡率データについて同様の計算ができる。この計算に用いられるデータの形式は非常に膨大なので本報に掲載することはできないが、より広範な原爆時年齢区分を用いた集計表(その結果は非常に類似している)を付録に掲載する。

これらのデータの正式な統計学的解析の多くは、実際の死亡日又は診断日とは無関係に、追跡調査全期間における癌発生率と線量レベルとの関係を示した。実際、これはある種の基本的な疑問に答えるのに

somewhat surprisingly (to these authors, at least) effective way of answering certain basic questions. However, understanding the temporal nature of the radiogenic excess is important, both from the theoretical viewpoint of studying radiation carcinogenesis, and from the more practical viewpoint of estimating lifetime risks for regulatory and medical decision-making. It is also important to carry out analyses using response times, in order to evaluate the adequacy of methods which ignore this aspect.

There have been informal investigations of temporal effects, but formal statistical methodology has rarely been used in these efforts. Some examples of reports focussing on temporal aspects of cancer incidence in the LSS are: Ichimaru et al,¹⁰ Ishimaru et al,¹¹ Land and Norman,¹² and Land and Tokunaga.¹³ The methods described in this paper have represented an effort to provide statistical tools for such investigation.

MATERIALS AND METHODS

Overview of the Data

The data analyzed here were reported by Ichimaru et al,¹⁰ and represent leukemia incidence during 1950-78 in the LSS extended sample. Incidence data was gathered from tumor registries, supplemented by reports of physicians treating those whose deaths were reported in the mortality study. Details of case definition and ascertainment are given in their report. The ascertainment is thought to be quite complete.

Only acute leukemia cases are used, since the temporal nature of chronic leukemia is believed to be different. The total number of cases used here is 117. Ichimaru et al¹⁰ report 135 acute leukemia cases; 13 of these are from a supplemental control sample of persons not in the cities ATB, which is not used here; 4 cases which occurred in persons of unknown dose were excluded. One more case occurred in persons aged over 60 ATB (exposed in Hiroshima, with a T65DR dose over 400 rad), a group excluded here for reasons given later. They also report 53 cases of chronic leukemia.

The radiation dose estimates used here (T65DR kerma) are for radiation incident on the body, taking into account external shielding but not shielding by body tissues. In some other leukemia analyses, estimates of bone marrow

幾分驚くべき(少なくとも本報の著者らにとっては)効果的な方法であると判明した。しかしながら、放射線誘発による過剰の経時的特性を理解することは、放射線誘発の癌発生の研究の理論的観点からしても、また、規定の医学的決定を下すための生涯リスクの推定という、より実際の観点からしても重要である。反応時間という側面を無視する技法の妥当性を評価するため、反応時間を用いた解析を行うことも重要である。

経時的効果の略式の研究はあるが、これらに正式な統計学的方法が用いられているのはまれである。寿命調査対象集団における癌発生率の経時的測面に焦点を当てた論文の例としては、市丸ら,¹⁰ 石丸ら,¹¹ Land 及び Norman,¹² 並びに Land 及び徳永¹³ がある。本報に述べた方法は、このような研究のための統計学的手段を提供する試みを示した。

材料及び方法

データの概観

本報で解析したデータは、市丸ら¹⁰ によって報告された寿命調査拡大集団における1950-78年の白血病発生率の資料を用いた。この白血病発生に関するデータは腫瘍登録から収集され、死亡調査に報告されている死亡者を治療した医師の報告も補足されている。症例の定義及び確認の詳細は彼らの論文に述べられている。症例の確認は極めて完全だと思われる。

慢性白血病の経時的特性は異なると思われるので、急性白血病の症例のみを用いた。本報で用いた総症例数は117である。市丸ら¹⁰ は急性白血病135例を報告しているが、このうち13例は原爆時市内不在者である対照集団からのものであり、本報では用いなかった。線量不明の4例も除いた。更に原爆時年齢60歳以上群に入る者(広島で被爆し T65DR 線量 400 rad 以上)が1例あったが、この群は後述の理由で本報では除外した。市丸らは慢性白血病も53例報告している。

本報で用いた放射線量推定値(T65DR kerma)は、外部遮蔽は考慮しているが体内組織による遮蔽は考慮していない身体への放射線入射量である。その他

dose have been used. Very roughly, these are one-half the dose estimates used here. The calculation of bone marrow dose involves separate estimates of gamma ray and neutron components, and of self-shielding factors, which are particularly uncertain in view of the current dosimetry reevaluation.

Table 1 gives, for each city, the distribution over dose categories of the number of persons initially in the study, the number of acute leukemia cases, and the number of person-years at risk during the follow-up period. As has been done in previous analyses of the LSS leukemia data,¹⁰ the approximately 27,000 LSS subjects who were not in the cities ATB are not included in this table or in subsequent analyses. The 2,386 subjects for whom detailed dose information is not available and all subjects aged over 60 ATB are also excluded. The period of study is October 1950 through December 1978. A person is considered to be at risk until developing leukemia or dying of any cause. The primary reason for studying the cities separately is that the radiation from the two bombs was qualitatively different.^{4,6}

の白血病の解析の中には、骨髄線量の推定値を用いているものもある。大ざっぱに言えば、その推定値は本報で用いた線量推定値の1/2である。骨髄線量の計算では、ガンマ線、中性子及び自己遮蔽因子を別々に推定するのであるが、これらの値は現在行われている線量再評価から考えると極めて不確実である。

表1に線量区分別の対象者数の分布、急性白血病症例数及び追跡期間中の観察人年数を都市別に表示した。寿命調査白血病データの前回の解析¹⁰で行われたように、原爆時市内不在者約27,000人の寿命調査対象者は、この表及び以降の解析には含まれていない。詳細な線量資料が得られない2,386人の対象者及び原爆時年齢60歳以上の全対象者も除外した。調査期間は1950年10月から1978年12月までである。対象者は白血病が発病するまで、又は何らかの原因で死亡するまで観察が必要であると考え、都市別に研究を行う主な理由は、両市の爆弾からの放射線が質的に異なるからである。^{4,6}

TABLE 1 ACUTE LEUKEMIA CASES AND PERSON-YEARS AT RISK AMONG A-BOMB SURVIVORS BY CITY AND DOSE, OCTOBER 1950 THROUGH DECEMBER 1978

表1 原爆被爆者における急性白血病症例及び観察人年数、都市及び線量別、1950年10月～1978年12月

	T65DR Dose in Rad							
	0	1-9	10-49	50-99	100-199	200-299	300-399	400+
Hiroshima								
Cases	30	10	12	6	11	3	7	11
Person-years at Risk	664770	385509	263201	66475	41294	15972	8793	12095
Number at Risk ^a	27575	15932	10908	2782	1739	659	368	510
Average Dose ^b	0	3	22	70	139	244	346	524
Nagasaki								
Cases	2	8	0	0	4	6	1	6
Person-years at Risk	98911	175013	97996	36072	35032	18279	6814	8889
Number at Risk	4000	7136	4030	1442	1388	722	270	374
Average Dose	0	4	21	71	147	244	344	529

^aNumber of people in the Life Span Study sample in October 1950.

1950年10月現在の寿命調査対象者数

^bDoses in excess of 600 rad are set equal to 600 rad in this calculation.

この計算においては、600 rad を越える線量は一律に600 rad とした。

The following type of dose-response analysis, using only data as presented in Table 1, is widely used.^{8,9} For a given city write y_d and R_d for the number of cases and time at risk, respectively, at dose level d . Consider fitting a model under which the y_d are independent Poisson observations with means $R_d\lambda_d$, where λ_d is the leukemia rate per person-year at risk, and is specified by a parametric model. Two models for λ_d particularly favored by radiation biologist, for the dose range here, are $\lambda_d = \beta_0 + \beta_1 d$ and $\lambda_d = \beta_0 + \beta_1 d + \beta_2 d^2$. Here d would represent class marks for dose categories. Models of this type are discussed in BEIR-3⁸ and NCRP-64.⁹ There is controversy over such models, centered substantially on the issue of estimation of risks at low levels of exposure (<10 rad), where the choice can make a large difference in the ratio of the estimated risks. The choice may be less critical for the purposes here, where focus is on modeling risks at higher doses, in order to estimate temporal effects.

A starting point for more detailed analysis is to consider the meaning of λ_d above. The risk of leukemia, both background ($d=0$) and radiogenic, is not constant but depends on, among other factors, age and time since exposure. The dependence on age ATB is well known, and that on time since exposure will be discussed below. The rate λ_d is a weighted average of these more specific rates, the time averaging being weighted by the survival distribution, considering all causes of death. While conceptually this is straightforward for a given age ATB, the fact that both the radiogenic excess and the survival fraction, at a given time since exposure, depend on age ATB makes the interpretation of λ_d complex.

Furthermore, the Poisson regression described above is the maximum likelihood procedure for censored data when the rates are modeled as constant in time.^{14,15} It will be seen that the rates are far from constant, so there should be serious concern about the optimality of this analysis.

Consideration of more specific rates raises the issue of differences by sex, as well. Sex is the only factor other than dose, age ATB, and time since exposure, which is generally available for the LSS subjects. The analysis here ignores the effects of sex. Although it would be of some

表1に示したデータのみを用いた次のような型の線量反応解析が広く用いられている。^{8,9} ある都市について線量レベル d での症例数、及び観察期間をそれぞれ y_d 及び R_d と表す。 y_d が、平均 $R_d \lambda_d$ の独立ポアソン観察値であるモデルが適合するとする。ただし、 λ_d は観察人年当たりの白血病発生率で、パラメトリックモデルで決定される。本報の線量範囲で放射線生物学者が特に支持する λ_d の二つのモデルは、 $\lambda_d = \beta_0 + \beta_1 d$ 及び $\lambda_d = \beta_0 + \beta_1 d + \beta_2 d^2$ である。ここで、 d は線量区分の階級値である。この型のモデルは BEIR-3⁸ 及び NCRP-64⁹ で論じられている。このようなモデルについては意見が分かれ、特に選択によって推定リスク率に大きな違いが生じ得る。低線量レベル (<10 rad) での被曝のリスク推定問題について議論が集中している。本報の目的のためには、この選択の問題はそれほど重要ではないと思われる。と言うのは、本報では経時的効果を推定するために、高線量でのリスクのモデルを作ることに焦点を絞っているからである。

より詳細な解析の最初に、上述の λ_d の意味を考える。バックグラウンド ($d=0$) リスクにせよ放射線誘発性リスクにせよ、白血病のリスクは一定ではなく、特に年齢及び被曝後の経過時間に依存する。原爆時年齢に依存することは周知のとおりであるが、被曝後の経過時間に対する依存について次に述べる。比率 λ_d は、これらのより特定化した比率の加重平均であり、全死因を考慮して生存率分布により重みをつけて時間の平均を求めた。概念的には、これはある原爆時年齢に対しては明瞭であるが、放射線誘発による過剰及び生存率の変動は、いずれも被曝後のある時点では原爆時年齢に依存するという事実により、 λ_d の解釈が複雑になる。

更に上述のポアソン回帰は、比率を時間によらず一定としてモデル化した場合の調査データの最大尤度法である。^{14,15} 比率は一定でないことが判明すると思われるので、この解析が最適であるかどうかについて注意深く検討しなくてはならない。

更に特定の比率について考えると、性による違いという問題も生ずる。性は、線量、原爆時年齢及び被曝後の経過時間以外で、寿命調査対象者について一般に得られる唯一の因子である。本報の解析は性の効果を見逃している。これを追究するのも興味深いことで

interest to pursue this, proper investigation of sex effects would have substantially complicated the presentation here. It is likely that the sorts of inferences considered here are largely "orthogonal" to effects of sex, but this has not been investigated thoroughly.

Table 2 illustrates, to an extent, the joint effects of age ATB and time since exposure on both the background rate and the radiogenic excess. This table contains incidence rates for both $d < 10$ rad and $d > 100$ rad, in a cross-classification of these two factors. Data from both cities are combined for this tabulation. Particularly notable is the strong suggestion that the temporal nature of radiogenic excess depends upon age ATB. This has been the conclusion of previous analyses, mentioned above. The aim here, in large part, is to provide significance tests and estimation schemes to study such apparent effects.

あろうが、もし性の効果を適切に研究したとすれば、この報告はかなり複雑なものになったであろう。本報で検討した推論は大部分性の効果に対して“直交”しているが、徹底的な研究は行われていない。

表2は、バックグラウンド比率及び放射線誘発による過剰に対する原爆時年齢と被曝後経過時間との同時効果のある程度示している。この表は、これら2因子の交差分類による $d < 10$ rad 及び $d > 100$ rad における発生率を示す。この表では両市のデータを組み合わせた。特に注目すべきことは、放射線誘発による過剰の経時的特性は、原爆時年齢に依存することが強く示唆されることである。これは上述の今回の解析からの結論である。本報の目的は、主にこのような明白な効果を研究するための有意性検定及び推定方法を提供することである。

TABLE 2 COMPARISON OF ACUTE LEUKEMIA INCIDENCE RATES BY TIME PERIOD AND TOTAL DOSE

表2 急性白血病発生率の比較、期間及び総線量別

Age at Time of Bomb	Dose (rad)	Time Period								
		Oct. 1950-Dec. 1958			Jan. 1959-Dec. 1966			Jan. 1967-Dec. 1978		
		Cases	PY ^a	Rate ^b	Cases	PY	Rate	Cases	PY	Rate
0-9	<10	1	89825	1.1	2	85796	2.3	2	127209	1.6
	100+	8	6681	120.	1	6350	15.7	0	9355	0.0
10-19	<10	2	97001	2.1	0	92091	0.0	8	134692	5.9
	100+	9	14546	61.9	2	13787	14.5	0	20000	0.0
20-34	<10	2	88663	2.3	0	83423	0.0	10	118702	8.4
	100+	4	10820	37.0	3	10143	29.6	4	14239	28.1
35-49	<10	6	101402	5.9	4	89303	4.5	11	108253	10.2
	100+	6	10884	55.1	3	9490	31.6	4	10966	36.5
50-60	<10	1	42058	2.4	0	30483	0.0	1	22875	4.4
	100+	1	3681	27.2	4	2558	156.	0	1716	0.0

^aPerson-years at risk during the time interval.

期間中の観察人年数

^bLeukemia incidence rate per 100,000 person-years at risk.

100,000観察人年数当たりの白血病発生率

Previous analyses of these data have discussed important questions and apparent effects. These issues cannot be addressed directly on the basis of the data as presented in Table 2 but will be discussed later. For example, some difference between cities in the temporal pattern of the radiogenic excess has been noted previously.¹⁰ This has been of particular interest due to the qualitative difference in radiation from the two bombs. Also, there is interest in any evidence that the induction period of radiogenic cancer depends on dose.

Formal statistical methods for such investigations are essential, because as the data are cross-classified into several dimensions, the estimation of correspondingly specific rates becomes less precise. Proper statistical methods should both provide efficient schemes for the investigation of potentially interesting interactions, and also help to avoid overinterpretation of apparent patterns in the data.

Organization of Calculations

It is useful to describe the general organization of calculations before turning to detailed formulation of models, as these issues are somewhat independent. Most of the calculations are done on grouped data, although some final results using ungrouped data are given for comparison. In a study involving about 80,000 persons, grouping greatly facilitates calculations, making it possible to investigate a variety of models through convenient interactive computing. In view of the complexity of the data, and uncertainty regarding appropriate models, this was considered to be very important.

The grouped data are numbers of leukemia cases, y , and person-years at risk, R , in a cross-classification by city, age ATB, dose, and time since exposure. The first three factors classify individuals and the fourth classifies their experience over time. The total time at risk, R , for a time interval is the sum over appropriate individuals of the elapsed time in that interval until they develop leukemia, or die of any cause, or the interval ends.

For age ATB the interval boundaries used are 0, 5, 10, 15, 20, 27, 35, 42, 50, and 60 years. Individuals aged over 60 ATB, among whom there has been only one acute leukemia case, are omitted here. This group was excluded since it is necessary to consider age ATB effects

これらのデータについて先般行った解析において重要な問題と明白な効果について検討したが、表2に提示したデータに基づいては直接検討できないので後段で論ずる。例えば放射線誘発による過剰の経時的パターンは、先に両市において差異が幾つか認められた。¹⁰ 二つの爆弾からの放射線の線質が異なるので、この点は特に興味深い。また、放射線誘発性癌の誘発時期は線量に依存するという証拠にも興味もたれる。

データは幾つかの次元に交差分類されるので、それに対応して特定の比率の推定値は正確ではなくなるので、このような調査には正式な統計学的方法が必要不可欠である。適当な統計学的方法は、興味深いものとなる可能性のある相互作用の究明のために効果的な方法を提供し、また、データの見掛け上のパターンの過大解釈を避ける一助となるであろう。

計算の構成

計算の一般的構成とモデルの詳しい形式とは幾分互いに独立しているので、後者に注目する前に前者について述べるのが有益である。計算は大部分、区分けしたデータを用いて行ったが、比較のために区分けしていないデータを用いた最終結果を幾つか示した。約80,000人に対する研究においては、区分けすることにより、便利な対話式演算法を用いて種々のモデルを検討することができ、計算がはるかに促進される。データが複雑なこと及び適当なモデルが確定していないことを考えると、これは非常に重要であると考えられた。

区分けしたデータは、白血病症例数 y 及び観察人年数 R を、都市、原爆時年齢、線量及び被曝後経過時間で交差分類したものである。最初の三つの因子は対象者を分類し、4番目の因子は対象者の経時的経験を分類する。ある期間区分の総観察期間 R は、その期間中、該当対象者が白血病を発病したり、何らかの原因で死亡したり、又は、その区間が終了するまでに経過した時間の合計である。

原爆時年齢については、区間境界は0, 5, 10, 15, 20, 27, 35, 42, 50及び60歳を用いた。原爆時年齢60歳以上の対象者に白血病が1例見られたが、ここではこの群は除外した。原爆時年齢が放射線誘発による過剰に及ぼす効果を考えることが必要であるが、

on the radiogenic excess, and this cannot be done reliably for this group because of the small number of person-years and cases. For the dose categories the boundaries are 0, 1, 10, 50, 100, 200, 300, and 400 rad. For follow-up time, the period October 1950 through December 1978 was divided into intervals of 2¼, 2, 2, 2, 4, 4, 6, and 6 years. Intervals of unequal length were chosen because of apparently more rapid temporal changes in excess risk during the early years. Thus there are (y, R) pairs for each of $2 \times 9 \times 8 \times 8 = 1,152$ cells. The tabulation in the Appendix differs only in that the 0-49 age range is divided into four rather than eight intervals.

For groupings by age ATB and dose, the approximation is simply to act as though all individuals are at class marks for the intervals. For age ATB, these are taken as midpoints. For dose, the class marks are taken as city-specific average doses within intervals; the values are given in Table 1. For this calculation, estimated individual doses over 600 rad were set to 600 rad, as is conventional at RERF. The rationale is that survival at such doses is unlikely enough to cast suspicion on the estimates.

The use of grouping on time of follow-up requires more elaboration; consider first the approach without that grouping. Let $i=1, 2, \dots$, denote cells, in any order, in the cross-classification of city, age ATB, and dose. Write $\lambda_i(t; \theta)$ for a given parametric model for the hazard function (risk) of leukemia incidence at time t since exposure, for those in cell i . Note that, given i , t also determines (approximately) age at risk, upon which the background rate, at least, will depend. The data can be thought of as a set of times to leukemia incidence or censoring t_{i1}, t_{i2}, \dots , for each i , along with corresponding indicators of whether t_{ij} is a censoring or incidence time. As noted previously, censoring in these data consists of either death from other causes or the end of the follow-up period under study. Denote by $\ell(\theta)$ the log likelihood for these data, under the usual assumptions about independent censoring (see for example, Kalbfleisch and Prentice¹⁴).

Now choose a partition of the follow-up period, and make the approximation that $\lambda_i(t; \theta)$ is constant in t within each interval of that partition. Denoting the intervals by j , the

この群は観察人年数及び症例数が少ないので、信頼できる考察を行うことはできない。したがってこの群は除外した。線量区分の境界は0, 1, 10, 50, 100, 200, 300及び400radである。追跡期間の1950年10月から1978年12月を2¼, 2, 2, 2, 4, 4, 6及び6年の間隔に分割した。初期には、過剰リスクの経時的変化が明らかに速いので不等間隔を選んだ。したがって $2 \times 9 \times 8 \times 8 = 1,152$ 個の各区分に(y, R)の組がある。付録表に0-49歳の年齢の幅が8区分ではなく4区分されている点のみが異なる。

原爆時年齢及び線量別の区分けについては、対象者はすべてその区間の階級値をとるという近似を行った。原爆時年齢では階級値は中央値とした。線量の場合、階級値は区間内での都市別平均線量とした。その値を表1に示す。この計算では、放影研の慣例どおり600rad以上の推定個人線量は600radとした。その理論的根拠は、そのような線量では生存はありえないので、推定値が問題になることはないということである。

追跡期間について区分けすることは更に説明を要する。まず、区分けをしない方法を考える。 $i=1, 2, \dots$, が順番にかかわらず都市、原爆時年齢及び線量を交差分類した区分を表すとする。区分*i*中の者について被曝後経過時間*t*における白血病発生率のハザード関数(リスク)の、あるパラメトリックモデルを $\lambda_i(t; \theta)$ とする。*i*が与えられれば、*t*は少なくともバックグラウンド比率が依存すると思われる観察年齢を(近似的に)決定することに留意されたい。このデータは、各*i*について t_{ij} が発見時期であるか、発病時期であるかを指示する添字とともに示された白血病発病時期と、発見時期 t_{i1}, t_{i2}, \dots , の組み合わせと考えられる。前述のように、これらのデータの検索は、他の原因による死亡又は追跡調査期間の終了から成る。独立検索(例えばKalbfleisch及びPrentice¹⁴を参照)に関する通常の仮定の下で $\ell(\theta)$ はデータの対数尤度を表す。

追跡期間の1区分を選び、 $\lambda_i(t; \theta)$ はその区分内の各区間内では*t*に対して一定であるという近似を行う。したがってその区間を*j*で表すと、結果として生ずる

resulting hazard function can thus be expressed as $\lambda_{ij}(\theta)$, $i, j=1, 2, \dots$. Write y_{ij} and R_{ij} for the corresponding number of cases and person-years at risk, respectively. For the moment consider the R_{ij} as fixed and the statistical model with c_{ij} as independent Poisson observations with means $R_{ij}\lambda_{ij}(\theta)$;

$$y_{ij} \underset{\text{ind}}{\sim} \text{Psn}\{R_{ij}\lambda_{ij}(\theta)\}, \quad i, j=1, 2, \dots \quad (1)$$

Although this is not in fact a valid model for the data, Holford¹⁵ has shown that the log likelihood for (1) is precisely the same as $\ell(\theta)$ described above. The key here is the assumption that $\lambda_i(t; \theta)$ is piecewise constant.

Strictly speaking, the model (1) is irrelevant, since the original log likelihood can be computed from the grouped data under the approximations discussed above. The authors find it quite useful to think of calculations in terms of (1), however, since such Poisson regression models are widely used and well understood. Also, for informal diagnostics of fit, it seems innocuous to think in terms of this model, with the R_{ij} as fixed. More formal diagnostics are based upon asymptotic theory for likelihood ratios.

Maximum likelihood estimation is done by the Newton-Raphson method, which can conveniently be thought of as iteratively reweighted (non-linear) least squares in model (1). Virtually no computational problems have been encountered, for dimensionality of θ up to 15. Tolerance of poor initial parameter values is provided by successive halving of the conventional Newton-Raphson step whenever the likelihood would otherwise decrease.

The rates $\lambda_{ij}(\theta)$ are functions of θ and the covariables; city, age ATB, dose, and time since exposure. The calculations for the grouped data are carried out interactively in GLIM¹⁶ with special macros. The procedure used by the analyst to specify models to be fit resembles that used for standard regression models in GLIM. Additional details of the organization of calculations are provided in the Appendix.

ハザード関数は $\lambda_{ij}(\theta)$, $i, j=1, 2, \dots$ と表される。対応する症例数及び観察人年数をそれぞれ y_{ij} 及び R_{ij} とする。差し当たり R_{ij} を固定したものと及び c_{ij} を用いた統計学的モデルを平均が $R_{ij}\lambda_{ij}(\theta)$ の独立ポアソン観察値と考えて

とした。実際には、これはデータに適したモデルではないが、Holford¹⁵ は(1)の対数尤度は上述の $\ell(\theta)$ と全く同じであることを示した。ここで重要な点は、部分的に見て $\lambda_i(t; \theta)$ が一定であるという仮定である。

厳密に言えばモデル(1)は不適當である。なぜなら、最初の対数尤度は、上述の近似の下で区分けしたデータから計算されるからである。しかしながら、我々は(1)を用いた計算を考えるのが非常に有益であると考える。それは、このようなポアソン回帰モデルが広く用いられており、よく理解されているためである。また、略式で適合度を判定する場合、 R_{ij} を固定してこのモデルで考えても差し支えないと思われる。より正式な判定は尤度比率の漸近理論に基づいて行う。

最大尤度推定は Newton-Raphson 法で行い、モデル(1)では簡便に反復加重した(非線形)最小二乗と考えられる。15までの次元数 θ については、実際に計算上の問題が起こったことはない。小さい初期パラメーター値の公差は尤度の減少の都度、従来の Newton-Raphson 法を逐次二等分する(尤度の減少を防ぐため)ことによって与えられる。

比率 $\lambda_{ij}(\theta)$ は θ と共変数、つまり都市、原爆時年齢、線量及び被曝後経過時間との関数である。区分けしたデータの計算は、特別マクロ命令を用いて GLIM¹⁶ で対話式に行われる。適合するモデルを指定するために解析者が用いる手法は、GLIM の標準回帰モデルに用いられる手法に類似している。計算の構成については付録で更に詳しく述べた。

Approach to Modeling

Approaches to the statistical modeling of cancer incidence in the LSS setting are discussed. First, attention is given to the general structure of such models, and then possible parametric formulations are described. The discussion in this section is not limited to models for leukemia incidence. The particular choices for the case of leukemia are presented later.

To simplify notation, models will be developed in terms of ungrouped data. Models for grouped data can be obtained from these as indicated before. The modeling of city differences will be discussed, but suppressed in the notation.

Define t =time since exposure, e =age ATB, d =dose level, and $a=e+t$ =age at time t . Write $\lambda(t;e, d)$ for the leukemia risk at time t for given (e, d) values.

A useful starting point for organizing thought is the simple model

$$\lambda(t;e, d) = \nu(a) + f(d)\rho(t), \quad (2)$$

where $f(d)$ increases from $f(0)=0$. Thus $\nu(a)$ is the age-specific background rate and $\rho(t)$ represents the temporal nature of radiogenic excess rate, modulated in a proportional hazards manner by the dose factor $f(d)$. Necessary enrichments of (2) will now be discussed, but the basic aim here is parametric estimation of functions similar to those involved in (2).

One necessary modification is to allow $\nu(a)$ to depend on e as well, to accommodate secular changes in age-specific background rates. The approach here is to assume tentatively that $\nu(a)$ can be replaced by a product of the form $\nu_1(e)\nu_2(a)$. Age ATB, e , is used here as a surrogate for birthdate. Further, possible differences between cities will be tentatively modeled with another multiplicative term. However, within the organization of computations used here, it is quite easy to allow for nonmultiplicative effects for these factors.

A major aspect of the data analysis in the LSS, is to determine whether, and how, the term $f(d)\rho(t)$ in (2) should be modified. Although this simple form seems to be of interest in

モデル化の方法

寿命調査対象集団における癌発生率を統計学的にモデル化する方法を論ずる。まず、このようなモデルの全体的構造に注目し、次に、可能なパラメーターの公式化について述べる。この章における考察は白血病発生率モデルに限定していない。白血病について特に選択したモデルは後述する。

表記を簡潔にするため、区分けしていないデータを用いてモデルを開発する。区分けしたデータを用いたモデルは、前述したように、区分けしていないデータから得られる。都市による差異のモデル化について述べるが、表記には表されない。

t =被曝後経過時間, e =原爆時年齢, d =線量レベル及び $a=e+t$ = t 時における年齢と定義する。ある (e, d) 値に対して t 時における白血病リスクを $\lambda(t;e, d)$ とする。

考察を組織化するための出発点としては次の単純なモデルが役立つ。

ただし $f(d)$ は $f(0)=0$ から増加する。したがって、 $\nu(a)$ は線量因子 $f(d)$ によって比例ハザードに修正された年齢別のバックグラウンド比率を表し、 $\rho(t)$ は放射線誘発による過剰率の経時的特性を表す。(2)に必要な強化を施すことについてはこれから論ずるが、ここでの基本的な目的は(2)に関するものと同様の関数のパラメトリック推定である。

年齢別のバックグラウンド比率の年次変化を調節するため、 $\nu(a)$ が e にも依存できるように修正することが必要である。ここでは、 $\nu(a)$ は $\nu_1(e)\nu_2(a)$ という形の積で置き換え可能であると暫定的に仮定するという方法をとる。出生日の代わりにここでは原爆時年齢 e を用いる。更に、都市による差異は暫定的に別の相乗項を用いてモデル化する。しかしながらここで用いた計算機構においては、これらの因子の非相乗効果を考慮することが極めて容易である。

寿命調査におけるデータ解析の主な側面は、(2)の項 $f(d)\rho(t)$ を修正するべきか否か、また、いかに修正すべきかを決定することである。理論上は、この単純な

principle, previous analyses and inspection of Table 2 strongly suggest that for leukemia, at least, $\rho(t)$ should also depend upon e . That is, those exposed at an early age seem to have a substantial radiogenic excess incidence sooner than those exposed at a later age. For other types of cancer, however, the apparent patterns are quite different.^{8,13} The possible effects of age ATB are indeed a major issue and require careful consideration. One relatively simple generalization is to replace $f(d)\rho(t)$ by a term of form $f(d)g(e)\rho(t)$. One interpretation of this term is that the level of excess, but not the temporal pattern, depends on e .

More general models of the form $f(d)\rho_e(t)$ must also be considered, where $\rho_e(t)$ denotes a class of functions of t , indexed by the variable e . Some simple such classes of functions are described later.

Another possibility of major interest, especially for cancers other than leukemia, is to replace $f(d)\rho(t)$ by $f(d)\nu(a)$, so that (2) becomes a full proportional hazards model.^{5,8} Note that this is within the very general form $f(d)\rho_e(t)$ mentioned above. In this case the partially nonparametric approach of Cox¹⁷ is an alternative to the fully parametric approach taken here. This will be discussed further.

Intercity differences in excess rates are tentatively modeled with the introduction of a multiplicative term, as was described for the background rate. This is attractive in view of the qualitatively different nature of the radiation, between cities, in that the city effect can be thought of as a factor applied to $f(d)$. A significance test relating to a more general difference between cities is described in a subsequent section.

Another possibility is that of more complicated interactions between d and t than can be represented by models of the form $f(d)\rho(t)$. For example, one concept of interest is often stated in terms of the induction period depending on dose. This might be modeled through the use of terms such as $f(d)\rho_d(t)$. This seems to be an attractive approach since it avoids elusive issues such as distributions of the random variable time-to-radiogenic cancer.

式が興味深いものであると思われるが、前回の解析及び表2の検討により、少なくとも白血病については $\rho(t)$ も e に依存することが強く示唆されている。すなわち若年で被曝した者には、より高齢で被曝した者より早い時期にかなりの放射線誘発による過剰発生が見られるようである。しかしながら、他の癌では見掛け上のパターンは全く異なる。^{8,13} 原爆時年齢の効果は実際重要な問題で、注意深く検討する必要がある。比較的単純な一般化の一つは $f(d)\rho(t)$ を $f(d)g(e)\rho(t)$ の項で置き換えることである。この項の解釈の一つは、経時的パターンではなく増加のレベルが e に依存するということである。

$f(d)\rho_e(t)$ のより一般的なモデルも検討しなくてはならない。ただし、 $\rho_e(t)$ は変数 e を添字とする t の類関数を表す。このような単純な類関数を幾つか後述する。

特に白血病以外の癌については、このほかに重要な可能性として $f(d)\rho(t)$ を $f(d)\nu(a)$ で置き換えることが考えられる。その場合(2)は完全な比例ハザードモデルになる。^{5,8} これは上述の非常に一般的な形 $f(d)\rho_e(t)$ の範囲内である点に注意されたい。この場合、ここで用いた完全パラメトリック手法に代えて、Cox の部分的ノンパラメトリック手法¹⁷ を用いてもよい。この点について更に論ずる。

過剰率の都市間の差異は、バックグラウンド比率で述べたように、相乗項を導入して暫定的にモデル化する。両市間の放射線質の違いを考えれば、都市の効果 $f(d)$ に適用される因子とみなし得るという点で、この方法は有益である。より一般的な都市間の差異に関する有意性検定については次章で述べる。

もう一つの可能性として、 $f(d)\rho(t)$ のモデルで表されるものよりも複雑な d と t の間の相互作用が挙げられる。例えば線量に依存する誘発期間という点から、しばしば、ある興味深い概念が述べられる。これは、 $f(d)\rho_d(t)$ のような項を用いることによってモデル化される。これにより確率変数時間対放射線誘発性癌の分布のような分かりにくい問題を回避することができるので、有益な手法と考えられる。

The generalization of (2) of primary interest here can be written as

$$\lambda(t; e, d) = \nu_1(e)\nu_2(a) + f(d)\rho_e(t). \quad (3)$$

This is an extremely general model, and will be severely restricted by the specific choice of parametric forms for the component functions. The aim now is to indicate some choices which have proved useful.

Experience suggests that the function $\nu_2(a)$ can be adequately represented, for many types of cancer, by taking its logarithm to be a function of $\log(a)$ which is linear, quadratic, or piecewise linear or quadratic with one or two change points. In the piecewise quadratic case it is useful to use splines, that is, to constrain not only the function but also its derivative to be continuous. It is known from studies of large populations^{18,19} that the log-log linear form takes on a central role in such studies. For many types of cancer, the linearity is remarkable over most of adult life, even when there is statistically significant curvature because of precise estimation. Often nonlinearities will be needed only to account for departures from this for the young (e.g., less than 25) or the elderly (e.g., greater than 70). In any case, the LSS data are adequate to provide checks of such assumptions in particular analyses.

If one assumes that the form $\nu_1(e)\nu_2(a)$ is suitable, it seems adequate ordinarily to take $\nu_1(e)$ to be of fairly simple form. For example, $\log\{\nu_1(e)\}$ may be taken as linear or quadratic in e , or constant on each of a choice of intervals for e . With the LSS data there is substantial confounding between estimation of $\nu_1(e)$ and $\nu_2(a)$. This is because the (current) follow-up period spans only 28 years, beginning at $e+5$ years. This confounding suggests that relatively simple models for $\nu_1(e)$ should be a goal, provided that attention is paid to diagnostics of fit.

Turning to the model for the radiogenic excess rate, the work reported here uses $f(d) \propto d$. Other possibilities are of interest, especially quadratic functions of d . It was felt that the linear approximation should be adequate for the purposes here, which primarily are to model the effect of moderately high doses in order to study the nature of the functions $\rho_e(t)$.

ここで最も興味深い(2)の一般化は

と表せる。これは極めて一般的なモデルであり、構成関数のパラメトリック形を特別に選択することで厳しく制限される。ここでの目的は、有益であると判明した選択を幾つか示すことである。

経験により、多くの型の癌について関数 $\nu_2(a)$ は、その対数を $\log(a)$ の線形関数、二乗関数、又は1~2点を修正すれば部分的線形又は二乗関数と考えることによって適切に表される。部分的二乗の場合スプラインを用いること、すなわち関数のみならず、その導関数も連続的であるようにすることが有益である。このような研究においては、対数-対数線形が中心的な役割を果たすことが、大集団を用いた研究^{18,19}から知られている。多くの型の癌について、正確な推定により統計的に有意な曲率が存在する場合でさえ、成人の寿命のほとんど全体にわたってこの線形性が顕著に認められる。しばしば若年者(例えば25歳未満)や、高齢者(例えば70歳以上)における線形からの逸脱を説明するためにのみ、非線形性が必要とされるであろう。いずれにせよ寿命調査データは、特定の解析におけるこのような仮定をチェックする上で適当である。

$\nu_1(e)\nu_2(a)$ という形が適当であると仮定すると、通常は $\nu_1(e)$ をかなり単純な形と考えるのが妥当と思われる。例えば $\log\{\nu_1(e)\}$ は e において線形又は二乗、あるいは e の各選択区間で一定と考えられることもある。寿命調査データにおいては、 $\nu_1(e)$ の推定及び $\nu_2(a)$ の推定でかなりの交絡がみられる。これは(現行の)追跡期間が、 $e+5$ 歳から始まって28年間しかないためである。この交絡から適合の判定に注意を払えば、 $\nu_1(e)$ に対しては比較的単純なモデルを目標にすべきことが示唆される。

ここで放射線誘発による過剰リスクについてのモデルを考えると、本報では $f(d) \propto d$ が用いられている。他の可能性、特に d の二乗関数にも興味をもたれる。ここでの目的は、主に関数 $\rho_e(t)$ の性質の研究のために中程度の高線量の効果をモデル化することであるが、そのためには線形近似が妥当であろうと思われた。

Choice of parametric models for the class of functions $\rho_e(t)$ is certainly the most difficult part of the analysis, since little is known a priori about this aspect, and the data are not adequate to estimate these functions with much precision. The general approach has been to select about five intervals for e , and allow the functions $\rho_e(t)$ to differ only between these intervals. For a given interval of e , models have been used in which $\log \{\rho_e(t)\}$ is a function of $\log(t)$ which is linear, quadratic, or piecewise linear or quadratic with one or two change points. At various stages of the analysis, differing degrees of richness in the models for $\rho_e(t)$ may be appropriate.

None of the assumptions discussed above are inherent to the general approach or computational scheme outlined here. The suggestions made are examples which the authors have found useful in analyzing several types of cancer incidence in the LSS.

To close this section, some discussion is in order regarding the decision to use parametric models, rather than the partially nonparametric approach that has developed out of the paper by Cox.¹⁷ One approach that has been considered at RERF starts from models of the form

$$\lambda(t; e, d) = \lambda_e(t)[1 + \beta d], \quad (4)$$

where $\lambda_e(t)$ is a class of arbitrary functions of t . Models of this type can be used by stratifying on e and proceeding as indicated by Kalbfleisch and Prentice (pp 87-89).¹⁴

For many types of cancer in the LSS, it appears that (4) must be generalized at least to the extent of allowing β to depend on e . Further, at least for leukemia, it would be important to include time-dependent covariables (op cit, pp 122-127)¹⁴ so that the ratio of radiogenic excess to background rate can also depend on t . Furthermore, it appears that this ratio needs to depend on e as well. The results of these generalizations of (4) are models which are as difficult to formulate as those developed from (2). If generalizations of (4) are used, the necessity of modeling the background rate is avoided. However, this is a relatively easy part of the parametric approach. On the other hand, what is lost in the propor-

類関数 $\rho_e(t)$ のパラメトリックモデルの選択という側面に関しては、ほとんど予備知識がない上、これらの関数を正確に推定できるほど十分なデータがないので、これは解析のうちで最も困難な部分である。一般には e について区間を約5個選択し、関数 $\rho_e(t)$ はこれらの区間の中だけで変化させる方法をとる。 e のある区間について、 $\log \{\rho_e(t)\}$ が $\log(t)$ の線形、二乗又は1ないし2個の修正点を含む部分的線形又は二乗関数であるようなモデルを用いた。解析の種々の段階で、異なる $\rho_e(t)$ のモデルの充実度を用いるのが適当であろう。

上述の仮定はいずれも、ここで概説した一般的方法や計算手法に固有のものではない。ここで行った提案は、著者らが寿命調査における数種の癌発生率の解析に役立つと考えたものである。

この章の終わりに当たり、Cox の論文¹⁷ から開発された部分的ノンパラメトリック法ではなく、パラメトリックモデルを用いるという決定を行ったことに関して、幾らか考察することが適切であろう。放影研で検討された一つの方法は

という形のモデルから出発する。ただし $\lambda_e(t)$ は t の任意類関数である。この型のモデルは e について階層を成し、Kalbfleisch 及び Prentice (87~89ページ)¹⁴ が指示する手続きを経ることによって用いることができる。

寿命調査における多くの型の癌について、少なくとも β が e に依存することを認める程度まで(4)を一般化しなくてはならない。更に少なくとも白血病については、バックグラウンド比率に対する放射線誘発による過剰の比も t に依存するように、時間依存的共変数(122-127ページ)¹⁴ を含めることが重要であろう。加えて、この比は e にも依存する必要があると思われる。(4)をこのように一般化した結果、モデルは(2)から作ったものと同じように定式化するのが困難なものとなった。(4)を一般化した場合、バックグラウンド比率をモデル化する必要はなくなる。しかしながら、

tional hazards approach is the ability to model the additive excess risk in a possibly simple manner. There are scientifically important questions which are difficult to express as generalizations of (4).

For cancers other than leukemia, proportional hazards models appear to be more viable. The authors have studied the suitability of these, within the parametric approach, by utilizing models of the form

$$\lambda(t; e, d) = \nu_1(e)\nu_2(a) [1 + f(d)\rho_e(t)]. \quad (5)$$

Parametric analyses for models of this type can be done similarly with the methods developed here; some details are given in the Appendix. Results for other sites will be given in later report.

RESULTS

Details of Model Fitting

This section provides detailed results of the application to the LSS leukemia incidence data, of the types of models described in the previous section, including discussion of the adequacy of fit. The model is similar to (3), with $f(d) = d$ and additional parameters, ξ_c and ψ_c , for intercity differences. Thus, the general form of the model is

$$\lambda(t; e, c, d) = \xi_c \nu_1(e)\nu_2(a) + d\psi_c \rho_e(t). \quad (6)$$

Consider now the choice of $\nu_1(\cdot)$, $\nu_2(\cdot)$, and the class of functions $\rho_e(\cdot)$.

All estimation is done by maximum likelihood from the grouped data except for one final result. Significance tests are based on the standard chi-square approximation to likelihood ratio tests. Approximate standard errors are obtained from the inverse information matrix for model (1). This, along with other aspects of computation, is discussed in the Appendix.

In many of the models to be discussed, separate parameters are estimated for each of five age ATB categories: 0-9, 10-19, 20-34, 35-49, and 50-60. For brevity these groups will be referred to as age cohorts. However, it should be remembered that the finer age ATB grouping described earlier is used for the computations.

これはパラメトリック法のうちでは比較的簡単な部分である。一方、比例ハザード法では、できるだけ簡単な方法で相加的過剰リスクをモデル化する能力が欠けている。(4)の一般化として表すのが困難な科学的に重要な問題がある。

白血病以外の癌については、比例ハザードモデルの方が実行可能と思われる。我々はパラメトリック法の範囲内で

という式のモデルを用いてこれらの適合性を研究した。この型のモデルのパラメトリック解析は、ここで開発された方法と同様の方法で行うことができる。詳細は付録に示した。他の部位癌についての結果は別に報告する。

結 果

モデル適合の詳細

この章では、寿命調査の白血病発生率データに前章で述べた型のモデルを適用した結果について詳説し、併せて適合の妥当性を論ずる。モデルは(3)と類似しており、 $f(d) = d$ 及び都市間の差異のための追加パラメーターとして ξ_c 及び ψ_c を有する。したがって、このモデルの一般式は

である。 $\nu_1(\cdot)$ 、 $\nu_2(\cdot)$ 及び関数 $\rho_e(\cdot)$ の選択を考える。

最終結果一つを除いて、推定はすべて区分けしたデータからの最大尤度によって行う。有意性検定は、尤度比率検定に対する標準 χ^2 近似に基づいたものである。近似的標準誤差はモデル(1)の逆情報行列から得る。これは演算の他の側面と併せて付録で論ずる。

考察するモデルの多くでは、原爆時年齢の5区分、0-9、10-19、20-34、35-49及び50-60のそれぞれに別々のパラメーターを推定する。簡潔にするためこれらの群を年齢コホートと呼称する。しかしながら、演算には前述のより細かい原爆時年齢区分を用いることを銘記すべきである。

The next few paragraphs deal with the selection of the functions $\nu_1(\cdot)$ and $\nu_2(\cdot)$ for the background rate. Following this, attention will be turned to choice of the class of functions $\rho_e(\cdot)$.

Data for subjects with $d < 10$ rad were used to investigate the nature of the background rate. It is believed that the radiogenic excess in these subjects may be considered negligible for this purpose.

Assuming only the multiplicative form $\nu_1(e)\nu_2(a)$, and ignoring intercity effects, crude background rates can be computed as follows. For a cross-classification on e and a of subjects with $d < 10$ rad fit the model

$$\log\{\nu_1(e_i)\nu_2(a_j)\} = \gamma_i + \delta_j,$$

where the γ_i and δ_j are free parameters which correspond to the intervals labeled by e_i and a_j . Parameter estimates and approximate standard errors are given below, for the indicated choice of intervals.

e_i	$\hat{\gamma}_i - \hat{\gamma}_1$	a_j	$\hat{\delta}_j$
0-9	0	0-19	-11.6 ± 1.0
10-19	-0.17 ± .67	20-29	-11.0 ± 0.6
20-34	-0.78 ± .81	30-39	-10.5 ± 0.6
35-49	-0.94 ± .88	40-49	- 9.5 ± 0.7
50-60	-2.6 ± 1.2	50-59	- 9.0 ± 0.9
		60-69	- 8.3 ± 0.9
		70+	- 8.1 ± 1.0

A plot of the $\hat{\delta}_j$ vs the logarithm of the age interval midpoints is almost exactly piecewise linear with a change point between the second and third points. Although this shape is not precisely estimated here, the Japanese national rates for acute leukemia as given by Otake²⁰ also suggest that such a piecewise linear model is reasonable. The change point of 27 years, used below, is approximately the maximum likelihood estimate from the data here.

以下の数段落では、バックグラウンド比率として関数 $\nu_1(\cdot)$ と $\nu_2(\cdot)$ との選択について述べる。これに続き類関数 $\rho_e(\cdot)$ の選択について検討する。

$d < 10$ rad の対象者のデータをバックグラウンド比率の性質を研究するのに用いた。この目的の場合、これらの対象者における放射線誘発による過剰は無視してもよいと思われる。

相乗式 $\nu_1(e)\nu_2(a)$ のみを仮定し、都市間の効果を無視すれば粗バックグラウンド比率は次のように計算される。 $d < 10$ rad の対象者の e 及び a を交差分類するとモデル

に適合する。ただし γ_i 及び δ_j は、 e_i 及び a_j で表した区間に該当する自由パラメーターである。本報で選択した区間のパラメーター推定値及び近似的標準誤差を以下に示す。

年齢区間の midpoint の対数に対して $\hat{\delta}_j$ を描くと、2番目の点と3番目の点の間を1点変更すれば、部分的にはほとんど正確に線形である。この形状はここでは正確に推定しないが、大竹の報告²⁰にあるように日本全国の急性白血病患者率からもこのような部分的線形モデルが妥当であることが示唆されている。下で用いた27歳という変更点は、ここでのデータからの近似的最大尤度推定値である。

A plot of the $\hat{\gamma}_i$ vs e indicates that a model linear in e with age-specific rates increasing in time is reasonable, considering the standard errors given. The trend on the age-specific rates is consistent with information presented in graphical form by Hirayama.²¹

This tentative background model is:

$$\log \{ \xi_c \nu_1(e) \nu_2(a) \} = \begin{cases} \log(\xi_c) + \theta_0 + \theta_1 e + \theta_2 \log(a), & a \leq 27 \\ \log(\xi_c) + \theta_0 + (\theta_2 - \theta_2^*) \log(27) + \theta_1 e + \theta_2^* \log(a) & a > 27 \end{cases} \quad (7)$$

When this model is fitted to the $d < 10$ rad data, the city effect ξ_c is not significant ($\hat{\xi}_c = -.02 \pm .38$) and this term is omitted henceforth.

A final check of the adequacy of the multiplicative form $\nu_1(e) \nu_2(a)$ was carried out as follows. The model (7) was fit to the $d < 10$ rad subjects, and the observed and fitted numbers of cases were calculated for a cross-classification over e and t . The intervals for e were taken as the five age cohorts used above, and the follow-up period was divided into five intervals of 4¼, 4, 4, 6, and 6 years. For each cell, a standardized residual, $(\text{observed} - \text{expected}) / \sqrt{\text{expected}}$, was computed. This residual is suggested by model (1). Informal inspection of this table indicated neither any systematic pattern of deviations, nor any individual cells of poor fit. Thus model (7) was judged to be adequate for the background rates.

Consider now the choice of parametric forms for the class of functions $\rho_e(\cdot)$. It seems that the approach to this major aspect of the analysis must focus on i) the use of appropriate significance tests to expose general patterns in the nature of the excess risk, ii) providing guidance regarding questions of interest which simply cannot be answered from the available data, iii) careful checking for systematic lack of fit to tentative models, and iv) avoidance of overfitting the data, that is, emphasizing what is probably only random variation in the data. As indicated earlier, the approach used is to group individuals into five age cohorts, and estimate a function $\rho_i(\cdot)$ for each group, $i = 1, \dots, 5$.

e に対して $\hat{\gamma}_i$ を描くと、標準誤差を考慮すれば、年齢別の比率が時間とともに上昇するような e に対して線形なモデルが妥当であると示唆される。年齢別の比率の傾向は平山²¹ のグラフに示される情報と一致する。

この暫定的バックグラウンドモデルは

である。このモデルを $d < 10$ rad のデータに適合させる場合、都市の効果 ξ_c は有意ではない ($\hat{\xi}_c = -.02 \pm .38$) ので、この項は以後省略する。

相乗式 $\nu_1(e) \nu_2(a)$ の妥当性の最終チェックは次のように行った。モデル(7)を $d < 10$ rad の対象者に適合させ、観察症例数及び適合症例数を e 及び t に関する交差分類について算出した。 e の区間は上記で用いた五つの年齢コホートをを用い、追跡期間は 4¼, 4, 4, 6 及び 6 年間の 5 区間に分けた。各区分ごとに標準化残差、 $(\text{観察値} - \text{期待値}) / \sqrt{\text{期待値}}$ を算出した。この残差は(1)から示唆される。この表を略式に検討したところ、いかなる逸脱の系統的パターンも、適合度の低い個々の区分も示されなかった。したがって、モデル(7)はバックグラウンド比率として妥当であると判断した。

関数 $\rho_e(\cdot)$ のパラメトリック式の選択について考える。解析のこの大きな側面への手掛かりとして、次のことに焦点を絞らなくてはならない。 i) 過剰リスクの性質の一般的パターンを明らかにするため、適当な有意性検定を行う。 ii) 利用可能なデータからは容易に答えることができない興味深い疑問に関して指針を与える。 iii) 暫定モデルへの適合の系統的欠如を注意深くチェックする。 iv) データの過剰適合を避ける。つまりデータのうち、確率変動の可能性のあるもののみを強調することを避ける。前述のように対象者を五つの年齢コホートに区分し、 $i = 1, \dots, 5$ の各群について関数 $\rho_i(\cdot)$ を推定する方法を用いた。

The choice of these functions is the most subtle part of the analysis. The data do not allow precise estimation of the functions $\rho_i(\cdot)$. This can be seen from the standard errors of estimates in the fitted models, and from likelihood ratio tests which compare alternative models. The fundamental reason for this is that the number of cases in each age cohort which are not predicted by the background model is small.

The number of excess cases for each cell (city, age ATB, dose, and time period group) for models such as (6) can be estimated as the product of person-years and the estimated value of the final term in (6). For the models described later in this section, the estimated numbers of excess cases for the five age cohorts are approximately 13(0-9), 12(10-19), 12(20-34), 15(35-49), and 5(50-60).

The form of the models chosen for $\rho_i(\cdot)$ will be presented here, along with some analysis to justify these choices. The course of research was to begin with somewhat more complicated models, which were eventually found to be unjustifiable. Whether one starts from simple models or more complex ones, it is at this stage that convenient interactive computing is particularly valuable.

The eventual choice of model for the radiogenic term in (6) was

$$\log\{\psi_c \rho_e(t)\} = \delta_c + \alpha_i + \beta_i \log(t), \quad (8)$$

where $c=0, 1$ denote Hiroshima and Nagasaki, respectively, and the subscript i is determined by the age cohort to which e belongs, $i=1, \dots, 5$. The model (6) with components (7) and (8) will be referred to as model (7-8).

Figure 1 illustrates the results of fitting model (7-8) to the grouped data. The dashed lines show, as a function of age, the estimated background rate for each of the five age cohorts. The solid lines show the estimated radiogenic excess rate in Hiroshima at 100 rad on the same scale. For clarity of presentation, smooth curves are sketched through points obtained by evaluating the model at class marks for the grouped data. The estimated excess rate in Nagasaki is 60% of that in Hiroshima. The parameter estimates for this and several other related models are given in Table 3.

これらの関数の選択は解析の中でも最も微妙な部分である。データから関数 $\rho_i(\cdot)$ を正確に推定することはできない。このことは適合させたモデルの推定値の標準誤差、及び代替モデルを比較する尤度比率の検定から分かる。この根本的理由は、各年齢コホートにおけるバックグラウンドモデルでは予測されない症例数が少ないからである。

(6)のようなモデルでの各区分(都市, 原爆時年齢, 線量及び期間群)ごとの過剰症例数は年及及び(6)の最終項の推定値との積として推定される。この章で後述するモデルの場合、五つの年齢コホートの推定過剰症例数は約13(0-9), 12(10-19), 12(20-34), 15(35-49)及び5(50-60)である。

$\rho_i(\cdot)$ のために選択されたモデルの式を、これらの選択の正当性を裏付ける解析と併せてここに示す。研究の方針としては、幾分より複雑なモデルから始めようとしたが、これらのモデルは結果的には妥当でないことが分かった。単純なモデルから出発しようと、より複雑なモデルから出発しようと、便利な対話式演算が特に貴重なのはこの段階である。

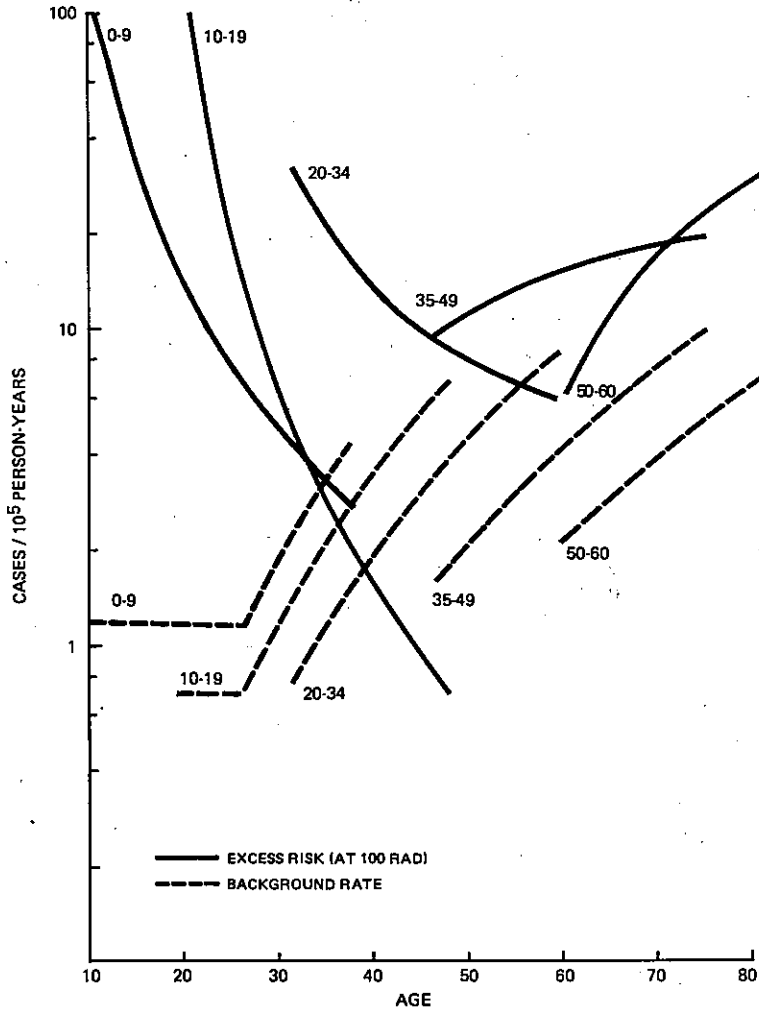
終局的には(6)の放射線誘発性の項のモデルとして

を選択した。ただし $c=0, 1$ はそれぞれ広島及び長崎を示し、添字 i は e が属する年齢コホートによって決定される ($i=1, \dots, 5$)。成分(7)及び(8)をもつモデル(6)はモデル(7-8)と呼称する。

図1に、区分けしたデータにモデル(7-8)を適合させた結果を示す。点線は、年齢の関数としての五つの年齢コホートのそれぞれについての推定バックグラウンド比率を示す。実線は、同尺度で広島100radでの推定放射線誘発による過剰率を示す。分かりやすく表すため、区分けしたデータの階級値でモデルを評価して得られた点を通るように滑らかな曲線を描いた。長崎の推定過剰率は広島のその60%であった。これらのパラメーター推定値及び幾つかの他の関連モデルを表3に示す。

FIGURE 1 BACKGROUND AND EXCESS (AT 100 RAD) ACUTE LEUKEMIA INCIDENCE RATES IN THE LIFE SPAN STUDY SAMPLE BY AGE AT EXPOSURE GROUP
 Separate Slopes for Radiogenic Excess (Model 7-8)

図1 寿命調査対象集団における急性白血病のバックグラウンド発生率及び(100radにおける)過剰発生率, 被爆時年齢群別放射線誘発による過剰の勾配(モデル7-8)



Diagnostics, discussed below, indicate that this model provides an adequate fit to the data. In view of the richness of the model, and the rather small number of apparently excess cases, there should be some concern about overfitting. However, this fit provides a basis for a number of significance tests of interest, and a reference point for assessing the adequacy of simpler models. No substantially simpler adequate model has been found.

以下で論ずる判定は、このモデルがデータに十分適合することを示唆している。モデルの充実度と、見掛け上の過剰症例数がかなり少ないことを考えると、過剰適合が懸念される。しかしながら、この適合により、多くの興味深い有意性検定の基礎が与えられ、より単純なモデルの妥当性を評価する上での参考になる。実質的により単純な妥当なモデルはまだ開発されていない。

TABLE 3 PARAMETER ESTIMATES AND STANDARD ERRORS FOR ACUTE LEUKEMIA INCIDENCE RATES IN THE LSS COHORT FOR SELECTED MODELS AND EXTENT OF GROUPING

表3 寿命調査コホートにおける急性白血病発生率のパラメーター推定値及び標準誤差、選択モデル及び区分けの程度別

Parameters ^a	Model/Grouping Used					
	Grouped Data					Ungrouped Data ^d
	Nine Age-at-Exposure Groups ^b			Five Age-at-Exposure Groups ^c		
	Unconstrained Model		Constrained Model			
Background	θ_0	-11.1 ±4.74	-11.0 ±4.64	-10.6 ±5.9	-11.6 ±4.84	
	θ_1	- 0.049±0.021	- 0.049±0.021	- 0.048±0.021	- 0.055±0.018	
	θ_2	0.078±1.48	- 0.030±1.45	- 0.15 ±1.83	0.20 ±1.52	
	θ_2^*	3.78 ±1.10	3.86 ±1.10	3.80 ±1.11	3.92 ±0.89	
Radiogenic Excess	δ	- 0.52 ±0.33	- 0.52 ±0.33	- 0.52 ±0.33	- 0.54 ±0.32	
	α_1	-14.0 ±0.57	-14.3 ±0.50	-14.2 ±0.50	-14.3 ±0.54	
	α_2	-15.0 ±0.80	-14.7 ±0.52	-14.7 ±0.51	-14.9 ±0.61	
	α_3	-13.9 ±0.45	-13.9 ±0.37	-13.9 ±0.37	-13.9 ±0.36	
	α_4	-13.4 ±0.33	-13.3 ±0.32	-13.3 ±0.32	-13.3 ±0.28	
	α_5	-13.0 ±0.55	-13.1 ±0.55	-13.0 ±0.52	-13.0 ±0.59	
	β_1	- 2.09 ±0.72	- 2.39 ±0.56	- 2.37 ±0.53	- 2.28 ±0.54	
	β_2	- 2.78 ±0.92				
	β_3	- 0.92 ±0.69	- 0.92 ±0.36	- 0.91 ±0.35	- 0.78 ±0.34	
	β_4	0.37 ±0.67	0.56 ±0.54	0.55 ±0.56	0.73 ±0.55	
	β_5	1.05 ±1.24				

^aThe background model is $\log \{ \nu(a, e, c) \} = \begin{cases} \theta_0 + \theta_1 e + \theta_2 \log(a) & a < 27 \\ \theta_0 + \theta_1 e + (\theta_2 - \theta_2^*) \log(27) + \theta_2^* \log(a) & a \geq 27 \end{cases}$
バックグラウンドモデルは

where a is attained age, e is age at exposure, and c is an indicator of city. The model for the unconstrained model for radiogenic excess (per rad) is $\log \{ \rho(t, e, c) \} = \delta c + \alpha_i + \beta_i \log(t/20)$, where $i=1, \dots, 5$ indexes the age cohorts (0-9, 10-19, 20-34, 35-49, 50-60, respectively), and t is time since exposure. For the constrained model, $\beta_1 = \beta_2 = \beta_3 - \gamma$ and $\beta_4 = \beta_5 = \beta_3 + \gamma$.

である。ただし、 a は到達年齢、 e は被曝時年齢、 c は都市を示す。(1 rad当たりの)放射線誘発による過剰の無条件モデルは $\log \{ \rho(t, e, c) \} = \delta c + \alpha_i + \beta_i \log(t/20)$ 、ただし $i=1, \dots, 5$ は年齢コホート(それぞれ0-9, 10-19, 20-34, 35-49, 50-60)を示し、 t は被曝後経過時間である。条件付きモデルでは、 $\beta_1 = \beta_2 = \beta_3 - \gamma$ 及び $\beta_4 = \beta_5 = \beta_3 + \gamma$ である。

^bThe nine age-at-exposure groups are 0-4, 5-9, 10-14, 15-19, 20-26, 27-34, 35-41, 42-49, and 50-60 years of age. The eight dose categories used are 0, 1-9, 10-49, 50-99, 100-199, 200-299, 300-399, 400+ rad. The eight time periods used are Oct. 1950-Dec. 1952, Jan. 1953-Dec. 1954, Jan. 1955-Dec. 1956, Jan. 1957-Dec. 1958, Jan. 1959-Dec. 1962, Jan. 1963-Dec. 1966, Jan. 1967-Dec. 1972, Jan. 1973-Dec. 1978.

九つの被曝時年齢群は0-4, 5-9, 10-14, 15-19, 20-26, 27-34, 35-41, 42-49及び50-60歳である。八つの線量区分は0, 1-9, 10-49, 50-99, 100-199, 200-299, 300-399, 400+radである。用いた八つの期間は1950年10月~1952年12月, 1953年1月~1954年12月, 1955年1月~1956年12月, 1957年1月~1958年12月, 1959年1月~1962年12月, 1963年1月~1966年12月, 1967年1月~1972年12月, 1973年1月~1978年12月である。

^cThe five age-at-exposure groups are 0-9, 10-19, 20-34, 35-49, 50-60 years of age. The groupings on dose and time period are the same as noted above. These data are presented in the Appendix.

五つの被曝時年齢群は0-9, 10-19, 20-34, 35-49及び50-60歳である。線量及び期間に関する区分けは上記と同様である。これらのデータは付録に示す。

^dAge-at-exposure and time-since-exposure were rounded to the nearest year.

被曝時年齢及び被曝後経過時間は最も近い年に丸めた。

Two questions related to the fit of this model are of particular interest; i) is it adequate to model the radiogenic excess as monotone, much less as log-linear, in t for each age cohort? and ii) is it reasonable to estimate the excess risk for the older age cohorts as increasing until the end of follow-up? For each age cohort in succession, a model was fit in which the corresponding term $\alpha_i + \beta_i \log(t)$ was replaced by a continuous function piecewise linear in $\log(t)$ with change points at $\log(13)$ and $\log(21)$, thus dividing the follow-up period into intervals of $8\frac{1}{4}$, 8, and 12 years. This involves, for each age cohort, two additional parameters. The likelihood ratio χ^2 statistics for the five cohorts are respectively 4.1, 1.4, 5.2, 1.6, and 7.9, each on 2 degrees of freedom (df). With the exception of the age 50-60 cohort, not only are these not significant, but the resulting estimates of the functions $\rho_i(t)$ are qualitatively similar to those shown in Figure 1. In particular, there is a sharp monotonic decrease in t for excess rates in the first two age cohorts. Excess rates in the third age cohort are not exactly monotone, but tend generally to decrease at a slower rate than for the first two age cohorts. In the fourth age cohort the rates tend generally to increase slowly in t , as in Figure 1.

For the age 50-60 cohort the fitted rates in the richer model are essentially zero for the first 6 and last 12 years of the follow-up period, and are somewhat higher between these points than those of Figure 1. This is because there were no cases at over 10 rad exposure during either the first $6\frac{1}{4}$ or last 12 years of follow-up. In addition there are only eight cases in total for this age cohort, and the fitted background model predicts about six cases. The most important issue is whether the lack of any cases in the final 12 years, for those at $d > 10$ rad, indicates that the true radiogenic excess rate has decreased to a markedly lower value than indicated in Figure 1. However, the expected number of excess cases for the final 12 years in the age 50-60 cohort computed from the model used for Figure 1 is only 1.5. If this model were true, the chance of seeing no radiogenic cases there would be $\exp(-1.5) = 0.2$. It seems clear that the temporal nature of radiogenic excess for this cohort is poorly estimated. The resolution taken here is to retain a model of the form (8) but somehow to combine the estimation of β_4 and β_5 .

このモデルの適合に関連する次の二つの疑問は特に興味深い。i) 放射線誘発による過剰を、各年齢コホートで t について単調と考えてモデル化することは妥当か、まして、対数線形と考えるのはどうか。ii) 高年齢コホートで過剰リスクが追跡期間の終わりまで増加し続けると推定することは正しいか否か。各年齢コホートに対して次々に対応する項 $\alpha_i + \beta_i \log(t)$ を $\log(13)$ 及び $\log(21)$ を変更点とする $\log(t)$ について部分的線形である連続関数と置き換えたモデルを適合させ、追跡期間を $8\frac{1}{4}$, 8 及び 12 年の区間に分割した。これには各年齢コホートについて追加パラメーターが二つ含まれ、五つのコホートについての尤度比率 χ^2 統計量はそれぞれ自由度 (df) 2 で 4.1, 1.4, 5.2, 1.6 及び 7.9 である。50-60 歳コホートを除いて、これらは有意でないだけでなく、結果として生ずる関数 $\rho_i(t)$ の推定値は図 1 に示すものと質的に類似している。特に初めの二つの年齢コホートでは、過剰率の t における急激な単調減少がみられる。第 3 番目の年齢コホートでの過剰率は厳密には単調でないが、初めの二つの年齢コホートよりも概してゆるやかに減少する傾向がある。第 4 番目の年齢コホートでは、過剰率は図 1 のように、概して t においてゆるやかに増加する傾向がある。

50-60 歳コホートでは、より充実したモデルの適合度は追跡期間の最初の 6 年間及び最後の 12 年間は本質的には 0 であり、図 1 の点よりもこれらの点の間の方が幾分高い。これは追跡期間の最初の $6\frac{1}{4}$ 年間にも最後の 12 年間にも、10 rad 以上の被曝症例がなかったためである。加えて、この年齢コホートには合計 8 例しかなく、適合させたバックグラウンドモデルでは約 6 例が予想されるのみである。最も重要な問題は、最後の 12 年間に $d > 10$ rad の症例がなかったことは、真の放射線誘発による過剰率が図 1 の値より著しく低い値にまで減少したことを示唆するかどうかである。しかしながら、50-60 歳のコホートの最後の 12 年間の推定過剰症例数を図 1 で用いたモデルから算出したところ、わずかに 1.5 である。このモデルが正しいとすれば、そこにおいて放射線誘発の症例が見られない確率は $\exp(-1.5) = 0.2$ であろう。このコホートの放射線誘発による過剰の経時的特性が不正確であることは明瞭であると思われる。ここでとられた解決策は、式(8)のモデルは維持するが、何らかの方法で β_4 及び β_5 の推定値を組み合わせることである。

The general issue of the extent of radiogenic excess at the end of follow-up, 1978 here, is important in consideration of plausible lifetime risks. This is not of such great interest for the age 50-60 cohort as for the others, since these persons are quite old by 1978. Of more importance is the plausibility of the results illustrated here that for the age 35-49 cohort the excess risk may still be substantial by the end of follow-up. One indication of this is that the fitted rates for the piecewise linear radiogenic excess in this cohort are not decreasing in the last time periods.

The strength of evidence that the β_i in model (7-8) are different is an important issue. Likelihood ratio tests indicate that no two β_i for adjacent age cohorts are significantly different, but the test that they are all equal gives $\chi^2 = 17.8$ on 4 df.

In view of the imprecise estimation of the β_i , and the generally increasing trend in them, it seemed appropriate to consider models with some smoothing of β_i . Modeling the β_i as linear in e results in rather extreme values of $\hat{\beta}_1$ and $\hat{\beta}_5$ which do not seem supported by the data and might be overemphasized in interpretation. Although the justification is rather weak, it was decided to impose the constraints $\beta_1 = \beta_2$ and $\beta_4 = \beta_5$, a hypothesis for which $\chi^2 = 0.8$ on 2 df. For this model, the three distinct slopes are almost exactly equally spaced. The imposition of this additional constraint has a totally negligible effect on the estimates, but reduces the standard error of β_3 somewhat. Thus, the model chosen for $\rho_e(t)$ can be written as

$$\log(\psi_c p_e(t)) = \delta c + \alpha_i + \beta \log(t) + \gamma I(e) \log(t), \quad (9)$$

where $i=1, 2, \dots, 5$ is an index for the age cohorts and $I(e)$ takes on the value -1 if $e < 20$, 0 if $20 \leq e < 35$, and 1 if $e \geq 35$. This model for the radiogenic excess together with background model (7) will be referred to as model (7-9).

Figure 2 shows the fitted cohort-specific background and radiogenic excess for this model, analogously to Figure 1. The parameter estimates for this model are given in Table 3.

追跡期間の終わり(ここでは1978年)における放射線誘発による過剰の程度という一般的な問題は、信頼できる生涯リスクを考えるのに重要である。50-60歳コホートの者は、1978年までにかなり高齢化しているので、この問題は他のコホートほど考慮する必要はない。より重要なことは、35-49歳コホートの過剰リスクは追跡期間の終わりにおいても、依然として高くなる可能性があるというここで示した結論が信頼できるかどうかという点である。これを示すもの一つに、このコホートの部分的線形の放射線誘発による過剰の適合度は、最後の期間で減少していないという事実が挙げられる。

モデル(7-8)の β_i が異なるという証拠が強力であることは重要な問題である。尤度比率検定で隣り合う年齢コホートのどの二つの β_i も有意差はないことが示されたが、 β_i はすべて等しいという検定では自由度4で $\chi^2 = 17.8$ である。

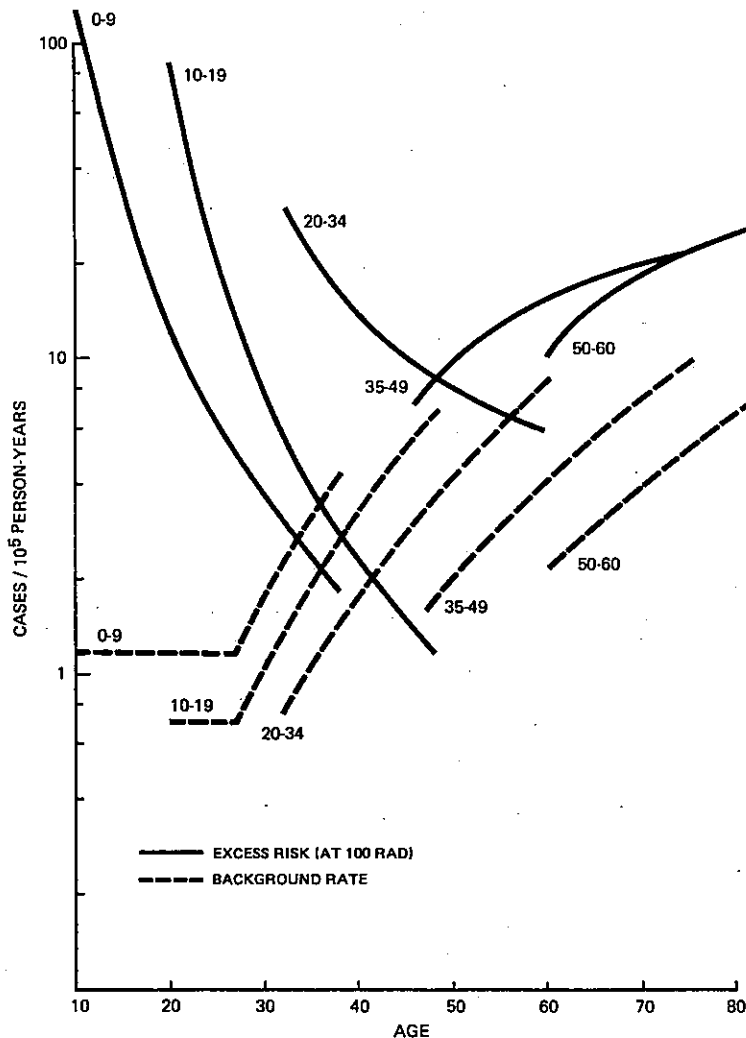
β_i の推定が正確でないこと、及び β_i が概して上昇傾向を示すことを考えれば、 β_i を幾分均一化してモデルを考えるのが妥当であると思われた。 β_i を e に対して線形としてモデル化すると、データによって裏付けられず、また、過剰解釈されることもある極端な $\hat{\beta}_1$ 及び $\hat{\beta}_5$ の値が得られるという結果になる。裏付けは若干弱い、自由度2で $\chi^2 = 0.8$ と仮定して $\beta_1 = \beta_2$ 及び $\beta_4 = \beta_5$ という制限を付けることにした。このモデルでは、3本の別々の曲線がほとんど厳密に等間隔に描かれる。この追加制限を付けたことによる推定値への影響は完全に無視できるが、 β_3 の標準誤差は若干減少する。したがって、 $\rho_e(t)$ のために選択されたモデルは

となる。ただし $i=1, 2, \dots, 5$ は年齢コホートを示し、 $I(e)$ の値は $e < 20$ ならば -1 、 $20 \leq e < 35$ ならば 0 及び $e \geq 35$ ならば 1 である。放射線誘発による過剰のこのモデルを、バックグラウンドモデル(7)とともにモデル(7-9)と呼ぶ。

図2は、図1に倣いこのモデルの適合させたコホート別バックグラウンド比率、及び放射線誘発による過剰を示す。このモデルのパラメーター推定値を表3に示す。

FIGURE 2 BACKGROUND AND EXCESS (AT 100 RAD) ACUTE LEUKEMIA INCIDENCE RATES IN THE LIFE SPAN STUDY SAMPLE BY AGE AT EXPOSURE GROUP
Constrained Model for Radiogenic Excess (Model 7-9)

図2 寿命調査対象集団における急性白血病のバックグラウンド発生率及び(100radにおける)過剰発生率, 被爆時年齢群別
放射線誘発による過剰の条件付きモデル(モデル7-9)



It is important to indicate in some way the level of uncertainty in estimation such as that shown in Figure 2. One such indication is obtained by computing the approximate standard error of $\hat{\alpha}_1 + \hat{\beta}_1 \log(t)$ for selected values of t . Table 4 gives the antilogs of these standard errors, based upon model (7-9), for times corresponding to the beginning, middle, and end of the follow-up period. These values can be interpreted as multiplicative standard errors for the excess rate.

図2に示すような推定値の不確か度を何らかの方法で示すことが重要である。これを示すものの一つは、選択された t 値について $\hat{\alpha}_1 + \hat{\beta}_1 \log(t)$ の近似的標準誤差の算出から得られる。表4に、モデル(7-9)に基づいて追跡期間の初期・中期・後期に該当する時期のこれらの標準誤差の逆対数を示した。これらの値は過剰率の相乗的標準誤差と解釈できる。

TABLE 4 MULTIPLICATIVE STANDARD ERRORS FOR RADIOGENIC EXCESS RATES AT SELECTED TIMES^a表4 選択した時点での放射線誘発による過剰率の相乗的標準誤差^a

Age at Time of Bomb	Years since Exposure					
	5		20		33	
	Hiroshima	Nagasaki	Hiroshima	Nagasaki	Hiroshima	Nagasaki
0-9	1.62 ^b	1.68	1.65	1.72	2.07	2.13
10-19	1.61	1.65	1.67	1.72	2.10	2.14
20-34	1.62	1.69	1.45	1.54	1.61	1.69
35-49	2.20	2.28	1.38	1.49	1.58	1.68
50-60	2.32	2.41	1.72	1.82	1.94	2.04

^aComputations are based upon model (7-9) which is described in the text.

演算は本文で述べたモデル(7-9)に基づく。

^bThese numbers are $\exp\{\hat{\sigma}_{\hat{y}_i(t)}\}$, where $\hat{y}_i = \hat{\delta}_c + \hat{\alpha}_i + \hat{\beta}_i \log(t)$ and $\hat{\sigma}_{\hat{y}_i(t)}$ is the approximate, asymptotic standard error of $\hat{y}_i(t)$. An approximate α -level confidence interval for $r_i(t) = e^{y_i(t)}$ is $[\hat{r}_i(t) / \exp\{\hat{\sigma}_{\hat{y}_i(t)} z_{\alpha/2}\}, \hat{r}_i(t) \exp\{\hat{\sigma}_{\hat{y}_i(t)} z_{\alpha/2}\}]$, where $z_{\alpha/2}$ is the $1 - \alpha/2$ percentile of the standard normal distribution.これらの数値は $\exp\{\hat{\sigma}_{\hat{y}_i(t)}\}$, ただし $\hat{y}_i = \hat{\delta}_c + \hat{\alpha}_i + \hat{\beta}_i \log(t)$ 及び $\hat{\sigma}_{\hat{y}_i(t)}$ は $\hat{y}_i(t)$ の近似的漸近誤差である。 $r_i(t) = e^{y_i(t)}$ の近似的 α レベルの信頼区間は $(\hat{r}_i(t) / \exp\{\hat{\sigma}_{\hat{y}_i(t)} z_{\alpha/2}\}, \hat{r}_i(t) \exp\{\hat{\sigma}_{\hat{y}_i(t)} z_{\alpha/2}\})$ ただし $z_{\alpha/2}$ は標準正規分布の $1 - \alpha/2$ 百分位数である。

The results presented in Table 5 are helpful in understanding the basis for the inferences drawn. This table presents, by age cohort and intervals of follow-up, the number of cases and the expected numbers of background and radiogenic excess cases based on the model (7-9). This table is helpful, for example, in seeing the justification for a model which provides estimates of excess risk monotonic in t for each age cohort. In particular, note that there is some indication of an increasing excess risk for the last 12 years of follow-up in the age 35-49 cohort.

It is of primary interest to use these models to estimate, and compare, the totals over time of radiogenic excess for each age cohort. The desired estimates are not those given by the integrated radiogenic risk, which would provide the total risk over time conditional upon remaining at risk for the entire follow-up. Rather, they must be computed by weighting the excess risk at time t by the chance of surviving all causes of death until that time. For purposes of illustration, the survival curves for these data by age cohort ignoring dose will be used. It should be noted that, except for cancer mortality, the survival rates of these A-bomb survivors is not different from that of the Japanese general population.

表5に示した結果は、導き出した推論の基礎を理解するのに役立つ。この表は、モデル(7-9)に基づいて年齢コホート及び追跡調査期間別に症例数、並びにバックグラウンド過剰症例及び放射線誘発性過剰症例の期待数を示したものである。この表は、例えば各年齢コホートについて、 t における単調な過剰リスクの推定値を与えるモデルの根拠を理解するのに役立つ。特に、35-49歳コホートに追跡期間の最後の12年間の過剰リスクの上昇が示唆されていることに留意されたい。

これらのモデルを、各年齢コホートについて放射線誘発による過剰の経時的合計の推定及び比較に用いることは非常に興味深い。望ましい推定値は、全観察期間中観察されることを条件として経時的リスク合計を与える総放射線誘発性リスクによって与えられるものではなく、むしろ t 時における過剰リスクを、そのときまでいかなる死因にもかかわらず生存している確率で加重して計算されなくてはならない。説明のため、これらのデータの線量を無視した年齢コホート別の生存曲線を用いる。癌死亡率を除いてこれらの原爆被爆者の生存率は、日本人一般の集団のそれと異なっていないことに留意すべきである。

TABLE 5 OBSERVED AND EXPECTED NUMBERS OF ACUTE LEUKEMIA CASES
BY AGE COHORT AND FOLLOW-UP PERIOD^a

表5 急性白血病症例の観察数及び期待数, 年齢コホート及び追跡期間別^a

Age at Time of Bomb	Cases	Time Period							Cohort Tables	
		Oct. 1950- Dec. 1952	Jan. 1953- Dec. 1954	Jan. 1955- Dec. 1956	Jan. 1957- Dec. 1958	Jan. 1959- Dec. 1962	Jan. 1963- Dec. 1966	Jan. 1967- Dec. 1972		Jan. 1973- Dec. 1978
0-9	Observed	3 (0) ^d	4 (0)	1 (1)	2 (0)	5 (1)	2 (1)	0	2 (2)	19 (5)
	Background ^b	0.43	0.38	0.38	0.37	0.74	0.73	1.45	2.95	7.43
	Excess ^c	4.9	2.38	1.46	0.98	1.18	0.69	0.60	0.36	12.59
10-19	Observed	6 (0)	1 (0)	2 (1)	3 (1)	1 (0)	2 (0)	5 (4)	4 (4)	24 (10)
	Background	0.28	0.25	0.26	0.30	0.78	1.24	3.09	5.24	11.42
	Excess	4.87	2.35	1.45	0.97	1.17	0.68	0.59	0.34	12.41
20-34	Observed	5 (2)	2 (0)	0	1 (0)	1 (0)	3 (0)	7 (3)	8 (7)	27 (12)
	Background	0.33	0.37	0.45	0.54	1.42	1.97	4.25	6.22	15.53
	Excess	2.14	1.49	1.23	1.05	1.71	1.37	1.62	1.26	11.87
35-49	Observed	1 (0)	1 (1)	2 (1)	8 (4)	2 (1)	6 (3)	6 (2)	13 (9)	39 (21)
	Background	0.74	0.75	0.86	0.97	2.31	2.84	5.21	6.11	19.87
	Excess	0.87	0.87	0.96	1.02	2.20	2.33	3.51	3.18	14.94
50-60	Observed	0	0	1 (1)	1 (0)	1 (0)	4 (0)	1 (1)	0	8 (2)
	Background	0.43	0.41	0.44	0.46	0.98	0.99	1.37	0.98	6.05
	Excess	0.43	0.42	0.44	0.45	0.90	0.83	0.95	0.57	4.98

^aExpectations are based upon model (7-9) which is described in the text.

期待数は本文で述べたモデル(7-9)に基づく。

^bExpected number of cases computed from person-years and background rates only.

人年及びバックグラウンド比率のみから算出した期待症例数。

^cExpected number of radiogenic cases computed from person-years, doses, and radiogenic rates only.

人年, 線量及び放射線誘発の比率のみから算出した放射線誘発による期待症例数。

^dNumbers in parentheses are observed numbers of cases among subjects with dose less than 10 rad.

括弧内の数値は線量10 rad以下の対象者の観察症例数である。

Since leukemia risk is being approximated as piecewise constant, the quantities to be estimated can be written as approximately

白血病のリスクは部分的に一定と近似されるので, 推定すべき量は下記のように近似的に表せる。

$$P_{ecd} = \sum_t d \rho_e(t) S_e(t) L_t \quad (10)$$

where the sum is over the eight time intervals, $S_e(t)$ is the survival probability at the beginning of the relevant interval for age cohort e , and L_t is the length of the interval. Note that (10) is linear in d . Although maximum likelihood estimates can be obtained simply by substituting parameter estimates, a modification was made as follows.

ただし, この合計は八つの時間区間の合計, $S_e(t)$ は年齢コホート e の該当区間の初期における生存確率, L_t は区間の長さである。(10)は d において線形であることに留意されたい。最大尤度推定値は単にパラメーター推定値を取り替えることで得られるが, 下記のように修正を行った。

Error in estimation of (10) can be assessed by the following bootstrap method.^{22,23} For each e and c , bootstrap values \tilde{P}_{ecd} of P_{ecd} were obtained by evaluating (10) at parameter values

(10)の推定上の誤差は次に述べるブートストラップ法^{22,23}で評価できる。各 e 及び c についての P_{ecd} のブートストラップ値 \tilde{P}_{ecd} は, パラメーター推定値及び

sampled from the multivariate normal distribution specified by the parameter estimates and their approximate covariance matrix. It was found that the resulting distribution of \tilde{P}_{ecd} are very nearly lognormal. The mean and standard deviations estimated for $\log(P_{ecd})$ in 200 trials were then transformed back to give estimates and multiplicative standard errors for P_{ecd} . These point estimates should have less bias than the maximum likelihood estimates.

The results are given in Table 6, as cases per 10^5 persons at $d=100$ rad. It is easily seen that there is no significant difference in excess risk between age cohorts, in spite of the different temporal nature of the risks.

その近似的共分散行列によって決定される多変量正規分布から抽出されたパラメーター値で(10)を評価することによって得た。結果として生ずる \tilde{P}_{ecd} の分布は対数正規に非常に近いことが分かった。200回の試行の $\log(P_{ecd})$ について推定した平均及び標準偏差は、 P_{ecd} の推定値及び相乗的標準誤差を求めるために元の形に変形した。これらの点推定値は最大尤度推定値よりも偏りが少ないはずである。

その結果を $d=100$ rad で100,000人当たりの症例数として表6に示した。リスクの経時的特性が異なるにもかかわらず、年齢コホート間で過剰リスクに有意差がみられないことが容易に分かる。

TABLE 6 BOOTSTRAP ESTIMATES^a OF EXCESS ACUTE LEUKEMIA INCIDENCE AT 100 RAD BY CITY AND AGE COHORT

表6 100radにおける急性白血病過剰発生率のブートストラップ推定値^a、都市及び年齢コホート別

City	Age at Time of Bomb				
	0-9	10-19	20-34	35-49	50-60
Hiroshima	449 ^b (1.38) ^c	286 (1.39)	301 (1.39)	388 (1.37)	306 (1.79)
Nagasaki	273 (1.45)	165 (1.41)	179 (1.52)	225 (1.49)	182 (1.84)

a—Based on 100 samples drawn from the approximate asymptotic distribution of the parameter estimates in model (7-9) which is described in the text.

本文で述べたモデル(7-9)のパラメーター推定値の近似的漸近分布から抽出した100例に基づく。

b—Excess cases per 100,000 persons with 100 rad T65DR dose during the period 1 October 1950 through 31 December 1978.

T65DRで100radの者10万人当たりの過剰症例数、1950年10月1日～1978年12月31日

c—Multiplicative standard error. An α -level confidence interval for an estimated excess, \hat{E} , with multiplicative standard error, m is $(\hat{E}/m^{z_{\alpha/2}}, \hat{E}m^{z_{\alpha/2}})$, where $z_{\alpha/2}$ is the $1-\alpha/2$ percentile of the standard normal distribution.

相乗的標準誤差。相乗的標準誤差 m の推定過剰数 \hat{E} の α レベルの信頼区間は $(\hat{E}/m^{z_{\alpha/2}}, \hat{E}m^{z_{\alpha/2}})$ である。

ただし $z_{\alpha/2}$ は標準正規分布の $1-\alpha/2$ 百分位数。

Two further analyses based upon model (7-9) were carried out to examine the consequences of the grouping used. The parameter estimates and standard errors for these fits are also given in Table 3. In one of these analyses, coarser grouping of age at exposure is used, these data are presented in the Appendix. For this analysis age cohort class marks were taken as 5, 15, 27, 42, and 55 years. This change is seen to have negligible effect on the parameter estimates and their standard errors. The other analysis uses data which are not grouped on any of the

使用した区分けの結果を調べるため、モデル(7-9)に基づく解析を更に二つ行った。これらの適合のパラメーター推定値及び標準誤差も表3に示した。これらの解析のうちの一つでは、より大きく区分けした被爆時年齢を用いた。このデータは付録に示した。この解析では年齢コホートの階級値は5, 15, 27, 42及び55歳とした。この変更が、パラメーター推定値及びその標準誤差に与える影響は無視できると思われる。もう一方の解析では、原爆時年齢、線量及び

variables; age ATB, dose, or time since exposure. Age and time are taken to the nearest year. The computational scheme for this was different from that outlined in the Appendix, consisting of rather straightforward, but computationally expensive, maximization of the likelihood.

Further Inference for the Leukemia Data

The examples presented in this section are concerned with inferences related to two specific areas of interest for the acute leukemia data, namely, city-by-time and dose-by-time interactions in models for excess risk. Since the particular questions of interest in the LSS data will vary markedly by cancer site, and since the specific inferences discussed may be more sensitive to the anticipated dose revisions, the purpose here is only to illustrate methodology for inferences which can be made after fitting models as described above.

Ichimaru et al¹⁰ have noted that excess leukemia may decline more rapidly in Nagasaki than in Hiroshima. Two methods of testing the significance of this result will be presented.

For a given age cohort, the model for excess risk suggested under Details of Model Fitting is:

$$d \exp \{ \delta c + \alpha + \beta \log(t) \} \quad (11)$$

where c is a 0, 1 indicator for city. The factor δc can be thought of as a city-specific scaling of dose. Indeed, the rationale for the city effect in this model is largely to allow for difference in the quality of the radiation between the two cities.

In the models used here there is some confounding between city-by-dose and city-by-time interactions in the radiogenic excess which allow a city-specific scaling of time of the form $e^{-\xi c t}$. That is, suppose that the integrated radiogenic risk in Nagasaki at $e^{-\xi t}$ is e^{δ} times that in Hiroshima at t . When this is done (11) becomes

$$d \exp \{ \delta c + \alpha + \beta \log(e^{-\xi c t}) \} e^{-\xi c} = d \exp \{ (\delta - \xi - \beta \xi) c + \alpha + \beta \log(t) \}. \quad (12)$$

被曝後経過時間のいずれの変数についても区分けしていないデータを用いる。年齢及び時間は最も近い年を選んだ。このための演算方法は付録で概説したものは異なり、かなり簡単ではあるが、演算費用の高い尤度の最大化から成る。

白血病データの更に詳しい推論

この章で述べる例は、急性白血病データで興味のもたれる二つの特定分野、すなわち、過剰リスクのモデルにおける時間×都市及び時間×線量の相互作用に関係する推論についてのものである。寿命調査データについて興味のもたれる特定の問題は癌の部位により著しく異なり、考察した特定の推論は予想される線量改訂により鋭敏に影響を受ける可能性があるため、ここでの目的は、上述したモデルの適合の後に可能な推論の方法を説明することのみである。

市丸ら¹⁰は白血病の過剰発生率は広島よりも長崎でより急速に減少することに気付いた。この結果得られた有意性検定のための二つの方法を示す。

ある年齢コホートについて、モデル適合の詳細の章で提案された過剰リスクのモデルは

である。ただし c は都市 (0, 1) を表す。因子 δ_c は線量の都市別尺度と考えられる。実際にこのモデルの都市の効果の理論的根拠は、主に両都市間の放射線質の違いを考慮することである。

ここで用いたモデルでは、 $e^{-\xi c t}$ の形で表される時間の都市別の尺度を認める放射線誘発による過剰は、都市×線量及び都市×時間の相互作用に幾らか交絡がみられる。すなわち、 $e^{-\xi t}$ における長崎での総放射線誘発性リスクは、 t における広島でのその e^{δ} 倍であると仮定すると、(11) は

Thus, within any collection of age cohorts among which β is constant, these forms of city-by-time and city-by-dose effects are confounded. This result is a consequence of the equivalence of scale changes and proportional changes in hazard functions for Weibull distributions. Kalbfleisch and Prentice (pp 34-35)¹⁴ have observed that this is true only for the Weibull distribution. However, it is likely that these interactions would be highly confounded in other models as well.

Since β varies with age cohort for acute leukemia, inference about both δ and ξ may be possible, provided ξ is not allowed to depend on cohort. Nevertheless, the joint effect of these two parameters is complex. In order to study these interactions, various models were fit and compared to a model analogous to (7-9) but with $\delta=0$. In the discussion below, likelihood ratio χ^2 statistics (LRS) relative to this model are examined for various models.

The addition of δ to the model described above leads to model (7-9), with a LRS of 2.94 on 1 df. As noted earlier, this is not statistically significant, but the estimated city-by-dose interaction is substantial (Nagasaki rate = .6 × Hiroshima rate). The addition of the time-scale factor to this model only provides an additional increase of .62 in the LRS, to 3.56 on 2 df. Not only is the estimated effect ($e^{-\xi} = 1.25$) not significant, but it is contrary to the effect suggested in earlier reports on these data. The inclusion of the dose-by-city interaction more than compensates for the apparent early decline in incidence rates in Nagasaki. This illustrates the complex nature of the joint effects of the two interactions.

If δ is held at zero and ξ is added, the LRS is .2 on 1 df. This suggests that whatever intercity differences exist in these data are better explained by a city-by-dose interaction than by a city-by-time-scale interaction.

Another way to model city-by-time interactions is to allow a city-specific shift in the age cohort slopes. Thus, β_e is replaced by $\beta_e + \xi c$. For this model, $\hat{\xi} = -1.1 \pm .77$, that is, slopes are smaller in Nagasaki, but not significantly so.

となる。したがって、 β が一定である年齢コホートのいかなる集合においても、これらの型の都市×時間及び都市×線量の効果は交絡する。この結果は、Weibull 分布のハザード関数における尺度変化と比例変化が同値であることから起こる。Kalbfleisch 及び Prentice (34-35ページ)¹⁴ は、これが Weibull 分布のみに当てはまることを発見した。しかしながら、これらの相互作用は他のモデルにおいても極めて交絡が起こりやすい。

急性白血病では β が年齢コホートとともに変化するるので、 ξ がコホートに依存しないようにすれば、 δ 及び ξ 双方についての推論が可能であるかもしれない。それにもかかわらず、これらの二つのパラメーターの同時効果は複雑である。これらの相互作用を研究するため、種々のモデルを(7-9)と類似のモデル(ただし $\delta=0$)に適合させ、比較した。下記の考察では、このモデルに関する尤度比率 χ^2 統計量(LRS)を種々のモデルについて調べた。

上述のモデルに δ を加えるとモデル(7-9)となり、LRS は自由度1で2.94である。既述のように、これは統計的に有意でないが、都市×線量の推定相互作用はかなり大きい(長崎の比率 = 0.6 × 広島 の比率)。このモデルに時間尺度因子を加えても LRS が更に0.62上昇して、自由度2で3.56となるだけである。この推定効果($e^{-\xi} = 1.25$)は有意でないだけでなく、これらのデータについてこれまでの報告で示唆された結果と反対である。線量×都市の相互作用を含めることにより、長崎での発生率の初期の明白な減少が十分に補われる。このことは、この二つの相互作用の同時効果の複雑な性質を示している。

δ を0にとどめ ξ を加えると、自由度1で LRS は0.2である。これは、これらのデータ中のどんな都市間の差異も都市×時間尺度の相互作用でよりも、都市×線量の相互作用でうまく説明されることを示唆している。

都市×時間の相互作用をモデル化する別の方法は、年齢コホート勾配に都市別の変動を認めることである。したがって、 β_e を $\beta_e + \xi c$ で置き換える。このモデルでは $\hat{\xi} = -1.1 \pm 0.77$ 、すなわち、勾配は長崎の方が

The LRS is 1.07 on 1 df, which is not significant. Note, however, that the estimate $\hat{\xi}$ corresponds to a substantial effect, which may be related roughly to what earlier investigators have noticed. There is an important lesson in its lack of significance.

Turning briefly to dose-by-time interactions an interesting model is the analog of (12)

$$d \exp \{ \delta c + \alpha + \beta \log(e^{\xi \log(d)} t) \} e^{\xi \log(d)} = d^{(1+\xi+\beta\xi)} \exp \{ \delta c + \alpha + \beta \log(t) \} \quad (13)$$

Thus, in age cohorts where β is constant, a scale change in t and the replacement of d by an estimated power of d are completely confounded. Over the dose range here, powers of d give models very similar to the popular quadratic dose-response models $\theta_1 d + \theta_2 d^2$. Again, for the models used here, the only hope of distinguishing between these two effects is given by variation of β between age cohorts. Assuming a common ξ for all age cohorts, the LRS for $\xi = 0$ is 1.11 on 1 df, which is not significant.

DISCUSSION

It seems that any method of analysis for these data which explicitly allows for temporal variation in risks is essentially equivalent to a Poisson regression of form (1). Within that framework, the only two possibilities seem to be i) completely parametric modeling, along lines discussed here, or ii) partially nonparametric modeling, using the partial likelihood method of Cox¹⁷ to eliminate some large subset of parameters in (1) before fitting. The results here suggest that approach i) is in fact viable. Approach ii) may be useful as well, especially when departures from proportional hazards can be modeled simply. Much remains to be learned about the relative merits of these two approaches.

Another related issue is whether it is necessary to model temporal variation in risks explicitly, in order to draw inferences about nontemporal matters such as the total risk over the follow-up period, or whether the dose response is linear in dose. This matter was touched upon in the Overview section in the discussion of inferences

小さいが有意ではない。自由度1でLRSは1.07であるが、これは有意ではない。しかしながら、推定値 $\hat{\xi}$ は実質的効果に相当することに注意すべきで、このことは初期の研究者たちの観察と概略的には関連している。このLRSが有意でないことは重要である。

線量×時間の相互作用に簡単に目を向けると、興味深いモデルは(12)に類似した

である。したがって、 β が一定である年齢コホートでは、 t の尺度変更と d を d の推定値での置き換えが完全に交絡している。ここでの線量範囲では、 d の冪により、一般的な二乗線量反応モデル $\theta_1 d + \theta_2 d^2$ に非常に類似したモデルが与えられる。ここで用いたモデルについても、これら二つの効果が区別される唯一の可能性は、年齢コホート間における β の変動により与えられる。全年齢コホートに共通の ξ を仮定すると、自由度1で $\xi = 0$ のLRSは1.11であるが、これは有意でない。

考 察

リスクの経時的変化を明白に考慮するどんなデータの解析方法も、本質的に(1)の形のポアソン回帰と同等であると思われる。その枠組み内でただ二つの可能性は、i)ここで論じた方針に沿う完全にパラメトリックなモデル化、あるいはii)適合させる前に(1)のパラメーターの大きな部分集合を除くための、Cox¹⁷の部分的尤度法を用いた部分的にノンパラメトリックなモデル化、であると思われる。ここでの結果から、i)の方法は実際に実行可能であると示唆された。ii)の方法も、比例ハザードからの逸脱が単純にモデル化できる場合には特に役立つ。これら二つの方法の相対的長所については多くこれからの研究が待たれる。

もう一つの関連問題は追跡期間中の全リスク、また、線量反応が線量に対して線形かどうかというような非経時的問題について推論するために、リスクの経時的変動を明白にモデル化することが必要か否かということである。この問題は表1からの推論の検討

based on Table 1. Further discussion is now given.

In the notation used here, some common analyses for dose response^{8,9} can be described as follows.

Let y_{ced} and R_{ced} denote the marginal totals, over time periods, of cases and person-years, respectively, for the cross-classification used in the analyses described above.

Let $\lambda_{ced}(\theta)$ be a parametric model for the risks per person-year over the entire follow-up period. Consider a Poisson regression, similar to (1),

$$y_{ced} \underset{\text{ind}}{\sim} \text{Psn}(R_{ced} \lambda_{ced}(\theta)). \quad (14)$$

The maximum likelihood estimates from (14) are precisely those for (1) if the temporal risk is taken to be constant in time, at the value $\lambda_{ced}(\hat{\theta})$.

Now consider estimation of the total risks over the follow-up period, as defined by (10). Even when the risks are not thought of as constant in time, these total risks can be estimated from model (14) as the product of $\lambda_{ced}(\hat{\theta})$ and the expected time at risk for age-group e . The question is whether such estimates are appropriate and optimal when the risk is not constant, in comparison with estimates based on model (7-9) which allows for nonconstant risks over time.

Table 7 contains bias-corrected estimates and multiplicative standard errors of the total risk over the follow-up period in Hiroshima by age cohort based on model (14) and, for comparison, the analogous bootstrap estimates derived from model (7-9).

This comparison indicates that (14) is a remarkably good model of the data for the present purposes, despite the marked changes in the hazard rates over time. Theoretical results to support this good approximation will be given in another report.

について概観の章で触れたが、ここでは更に詳しく考察する。

一般的な線量反応の解析^{8,9}は、ここで用いた表記で次のように示される。

y_{ced} 及び R_{ced} はそれぞれ、上述の解析で用いた交差分類における症例数・人年数の経時的周辺合計を表すとする。

$\lambda_{ced}(\theta)$ は全追跡期間中の人年当たりのリスクのパラメトリックモデルとする。(1)と同様のポアソン回帰

を考える。(14)からの最大尤度推定値は、経時的リスクが時間に対して一定とすれば、 $\lambda_{ced}(\hat{\theta})$ という値で、正確に(1)の最大尤度推定値である。

(10)で定義されるような追跡期間中の全リスクの推定について考える。リスクを時間に対して一定と考えない場合でも、これらの全リスクは $\lambda_{ced}(\hat{\theta})$ と年齢群 e の期待観察時間との積としてモデル(14)から推定できる。問題はリスクが一定でない場合に、このような推定値はリスクが経時的に一定でないことを考慮したモデル(7-9)に基づく推定値と比較して、妥当かつ最適なものであるか否かである。

表7に、モデル(14)に基づく年齢コホート別の広島での追跡期間中の全リスクの偏りを修正した推定値、及び相乗的標準誤差を示すとともに、比較のためモデル(7-9)から導いた類似のブートストラップ推定値を示した。

この比較により、ハザード率が経時的に著しく変化するにもかかわらず、(14)はここでの目的には著しく良いデータモデルであることが示唆された。この良い近似を支持する理論上の結果は別の報告で発表する。

TABLE 7 COMPARISON OF EXCESS ACUTE LEUKEMIA INCIDENCE ESTIMATES FOR HIROSHIMA BY AGE COHORT AT 100 RAD

表7 広島100radにおける急性白血病過剰発生率推定値の比較, 年齢コホート別

Model	Age at Time of Bomb				
	0-9	10-19	20-34	35-50	50-60
Constant Risk ^a	392 (1.44) ^b	236 (1.46)	305 (1.41)	392 (1.34)	281 (1.64)
Time-dependent Risk ^c	449 (1.38)	286 (1.39)	301 (1.39)	388 (1.37)	306 (1.79)

a— Bias-corrected estimated excess cases per 100,000 persons with 100 rad T65 DR dose in Hiroshima during the period 1 October 1950 through 31 December 1978. Estimates are based on the model $\lambda_{edc}(t) = \exp\{\nu_0\} + d \exp\{\rho_e + \mu c\}$. The estimate is $100 r_e \exp\{\hat{\rho}_e - \frac{1}{2} \hat{\sigma}^2 \hat{\rho}_e\}$, where r_e is the total person-years at risk for age cohort e and $\hat{\sigma}^2 \hat{\rho}_e$ is the estimated standard error of ρ_e .
 1950年10月1日から1978年12月31日までの期間中の広島におけるT65DRで、100radの者10万人当たりの偏りを訂正した推定過剰症例数。推定値はモデル $\lambda_{edc}(t) = \exp\{\nu_0\} + d \exp\{\rho_e + \mu c\}$ に基づく。その推定値は $100 r_e \exp\{\hat{\rho}_e - \frac{1}{2} \hat{\sigma}^2 \hat{\rho}_e\}$ である。ただし r_e は、年齢コホート e の総観察人年数で、 $\hat{\sigma}^2 \hat{\rho}_e$ は $\hat{\rho}_e$ の推定標準誤差である。

b— Multiplicative standard error of the estimated excess.
 推定過剰数の相乗的標準誤差。

c— Bootstrap estimate of excess based on model (7-9).
 モデル(7-9)に基づくブートストラップ法による推定過剰数。

APPENDIX

Estimation in (1) is a straightforward Poisson regression if the rates, λ , are log-linear in the unknown parameters. However, this is not the case for the type of models (2-4) developed in the text. It is possible, however, to formulate general classes of models to facilitate computing. The results for grouped data presented earlier were obtained using a convenient interactive procedure developed in GLIM.¹⁶ This Appendix contains a general description of the class of models used and an outline of the GLIM procedure. The grouped acute leukemia incidence data are presented in Appendix Table 1.

Let Y and R be GLIM variables containing numbers of cases and person-years at risk for the cross-classification used here. Denote by x_j , $j=1, \dots, p$, further GLIM variables which can be calculated as any desired functions of the class marks. Letting λ_i denote the rate in the i th cell of the cross-classification, and d denote dose, two useful classes of models are given by:

$$\lambda_i(x_1, \dots, x_p) = \exp\left\{\nu_0 + \sum_{i=1}^k \nu_i x_i\right\} + d \exp\left\{\rho_0 + \sum_{i=k+1}^p \rho_{i-k} x_i\right\}$$

and 及び

$$\lambda_i(x_1, \dots, x_p) = \exp\left\{\nu_0 + \sum_{i=1}^k \nu_i x_i\right\} [1 + d \exp\left\{\rho_0 + \sum_{i=k+1}^p \rho_{i-k} x_i\right\}].$$

付 録

(1)による推定は、比率 λ が不明のパラメーターについて対数線形の場合は、簡単なポアソン回帰である。しかしながら、本文で開発された型のモデル(2-4)については、これは当てはまらない。しかしながら、演算を促進するためにモデルの一般的な階級を定式化することは可能である。先に述べた区分けしたデータの結果は、GLIM¹⁶で開発された便利な対話式手法を用いて得られた。この付録では、用いたモデルの階級及びGLIMの手法を概説する。区分けした急性白血病発生率データを付録表1に示す。

Y 及び R を、ここで用いた交差分類による症例数及び観察人年数を含むGLIM変数とする。階級値のいかなる関数としても算出される他のGLIM変数を $x_j (j=1, \dots, p)$ で表す。 λ_i は交差分類の i 番目の区分の比率、 d は線量を表すとすると、二つの有益なモデルの階級は

These models might be called generalized additive and multiplicative models, respectively.

To fit a given model, the user simply calculates desired variables x_1, \dots, x_p and invokes a macro for the fitting. Rather than having the user specify k and p , it was convenient to fix them in advance and require the user to set any unused x -variables to zero.

Fitting is iterative and requires initial values for parameters. At the outset, it is acceptable to set ν_0 and p_0 to reasonable values and the other parameters to zero. In fitting a succession of models, estimates from previous fits will often work well.

Calculations can be organized as follows. Write Y , R , and $\lambda(\theta)$ for vectors of cases, person-years, and rates given by a model such as above. Let $D(\theta)$ denote the matrix $\partial\lambda/\partial\theta$, so that the linear approximation at a value θ_0 in a sequence of iterations is $\lambda(\theta) \doteq \lambda(\theta_0) + D(\theta_0)(\theta - \theta_0)$. Write $R\lambda$ for the vector with elements $R_i\lambda_i$.

Apart from numerous bookkeeping activities to make user interactions easier, the main function of the GLIM procedure is to compute successively the matrix $D(\theta_0)$, and fit by ordinary weighted least squares the model $Y - R\lambda(\theta_0) = D(\theta_0)(\theta - \theta_0) + e$, $\text{Var}(e) = \text{Diagonal}\{R_i\lambda_i(\theta_0)\}$. The actual likelihood is easily computed in terms of model (1), this can be used to terminate iterations and to provide for likelihood ratio tests. The covariance matrix of estimates $(\theta - \theta_0)$ routinely provided by GLIM for the weighted least squares regression just described is a suitable approximation to standard errors of estimates. This is $[D(\hat{\theta})' \text{Diagonal}\{R_i\lambda_i(\hat{\theta})\}^{-1} D(\hat{\theta})]^{-1}$, which is the inverse of the (expected) information matrix for model (1). A listing of the GLIM macros is available from RERF.

Appendix Table 1 contains the grouped acute leukemia incidence data for the LSS from the period 1 October 1950 through 31 December 1978. This data is identical to that used for the modeling discussed in the text except for the age ATB grouping. In this tabulation the five age ATB groups (0-9, 10-19, 20-34, 35-49, and 50-60) correspond to the age cohorts. The analyses in the text are based on nine age ATB groups (0-4, 5-9, 10-14, 15-19, 20-26, 27-34, 35-41, 42-49,

と与えられる。これらのモデルはそれぞれ、一般相加的及び一般相乗的モデルと呼ぶこともできる。

あるモデルを適合させるためには、ユーザーは単に変数 x_1, \dots, x_p を算出し、適合のためのマクロ命令を呼び出すだけでよい。ユーザーが k 及び p を指定するのではなく、前もってそれらを固定しておき、ユーザーは使用していない x 変数を 0 に設定する方が便利であった。

適合させることは反復的であり、パラメーターの初期値を必要とする。最初に ν_0 及び p_0 を適当な値に、その他のパラメーターを 0 に設定することは認められる。一連のモデルを適合させる際に、前回の適合から求めた推定値がしばしばうまく働く。

計算は次のように構成される。上述のようなモデルにより与えられる症例数、人年数及び比率のベクトルを Y , R 及び $\lambda(\theta)$ とする。 $D(\theta)$ が行列 $\partial\lambda/\partial\theta$ を表すとすれば、反復系列中の値の θ_0 での線形近似は $\lambda(\theta) \doteq \lambda(\theta_0) + D(\theta_0)(\theta - \theta_0)$ である。 $R_i\lambda_i$ を要素にもつベクトルを $R\lambda$ とする。

ユーザーのインタラクションをより簡単にするための膨大な簿記作業とは別に、GLIM の手法の主要機能は、行列 $D(\theta_0)$ を連続的に演算を行い、普通に加重された最小二乗数により、モデル $Y - R\lambda(\theta_0) = D(\theta_0)(\theta - \theta_0) + e$, $\text{Var}(e) = \text{Diagonal}\{R_i\lambda_i(\theta_0)\}$ を適合させることである。実際の尤度はモデル(1)により簡単に演算が行われ、これは反復を終了させ尤度比率検定を行うために用いることができる。今述べた加重最小二乗回帰に対してルーティンに GLIM で与えられる推定値 $(\theta - \theta_0)$ の共分散行列は、推定値の標準誤差への妥当な近似である。これは $[D(\hat{\theta})' \text{Diagonal}\{R_i\lambda_i(\hat{\theta})\}^{-1} D(\hat{\theta})]^{-1}$ で、モデル(1)の(期待)情報行列の逆行列である。GLIM マクロ命令のリストは放影研で入手できる。

付録表 1 に、1950年10月1日から1978年12月31日までの期間の寿命調査データの区分けした急性白血球の発生率を示す。このデータは原爆時年齢による区分けを除いて、本文中で論じたモデル化に用いたものと同一である。この表において、五つの原爆時年齢群 (0-9, 10-19, 20-34, 35-49 及び 50-60) が年齢コホートに相当する。本文の解析は、九つの原爆時年齢群 (0-4, 5-9, 10-14, 15-19, 20-26, 27-34, 35-41, 42-49 及び 50-60) に

and 50-60). The parameter estimates obtained when model (7-9) is fit to the condensed data presented here are given in Table 3.

基づいている。モデル(7-9)をここに挙げた凝縮データに適合させた場合に得られたパラメーター推定値を表3に示す。

APPENDIX
付録

TABLE 1 ACUTE LEUKEMIA INCIDENCE DATA 1 OCTOBER 1950 - 31 DECEMBER 1978
表1 急性白血病発生データ, 1950年10月1日-1978年12月31日

Age at Exposure	Dose	Time Period															
		Oct. 1950- Dec. 1952		Jan. 1953- Dec. 1954		Jan. 1955- Dec. 1956		Jan. 1957- Dec. 1958		Jan. 1959- Dec. 1962		Jan. 1963- Dec. 1966		Jan. 1967- Dec. 1972		Jan. 1973- Dec. 1978	
		C	PY	C	PY	C	PY	C	PY	C	PY	C	PY	C	PY	C	PY
Hiroshima																	
0-9	0	0	11736.00	0	10385.64	1	10390.45	0	10357.74	0	20657.59	0	20573.94	0	30718.61	0	30470.30
	1-9	0	6679.54	0	5910.71	0	5909.97	0	5892.49	0	11747.69	0	11713.93	0	17476.56	1	17287.47
	10-49	0	4587.74	1	4057.01	0	4053.02	0	4039.69	2	8068.25	1	8042.30	0	12023.61	0	11937.64
	50-99	0	1026.55	0	907.38	0	908.16	0	904.83	1	1800.32	0	1793.99	0	2680.13	0	2662.15
	100-199	2	508.15	0	447.18	0	444.30	0	443.70	0	888.00	0	884.04	0	1317.46	0	1307.79
	200-299	0	225.46	0	199.86	0	200.14	0	199.73	0	390.10	0	384.00	0	573.91	0	550.76
	300-399	0	99.20	0	87.94	0	84.96	0	83.94	0	168.00	0	168.00	0	250.59	0	245.78
400+	0	130.77	1	113.16	0	112.08	0	111.92	1	220.95	0	218.04	0	321.44	0	317.71	
10-19	0	0	11991.00	0	10561.88	0	10501.05	0	10411.35	0	20685.99	0	20546.45	4	30542.15	3	29983.53
	1-9	0	7302.87	0	6450.85	0	6432.29	0	6397.91	0	12724.34	0	12597.83	0	18703.47	1	18347.91
	10-49	0	4244.15	0	3752.79	0	3736.69	0	3712.05	0	7373.27	1	7306.15	1	10824.46	0	10605.57
	50-99	1	1021.45	0	900.34	0	898.35	0	894.26	0	1782.27	0	1760.59	0	2618.06	0	2568.98
	100-199	0	820.68	0	723.72	0	720.49	2	713.39	1	1407.72	0	1392.39	0	2052.80	0	2005.54
	200-299	0	334.55	1	295.16	0	292.27	0	289.67	0	573.20	0	568.32	0	839.43	0	816.69
	300-399	0	241.31	0	211.85	0	212.15	0	210.39	0	420.00	0	417.79	0	620.62	0	594.33
400+	2	295.10	0	259.82	0	260.18	0	259.82	0	514.19	0	508.00	0	739.87	0	689.64	
20-37	0	2	12198.10	0	10727.48	0	10681.86	0	10594.68	0	21025.06	0	20750.25	2	30409.27	4	29135.06
	1-9	0	7339.46	0	6455.09	0	6413.29	0	6362.92	0	12600.08	0	12375.68	0	18103.89	2	17471.44
	10-49	0	4916.37	0	4316.29	0	4287.92	1	4252.73	1	8431.56	0	8302.23	0	12133.50	0	11643.25
	50-99	0	1365.72	1	1200.49	0	1196.82	0	1185.44	0	2345.65	0	2300.14	0	3359.75	1	3194.58
	100-199	0	812.09	0	710.21	0	707.81	0	702.79	0	1382.95	1	1357.36	2	2001.99	0	1909.47
	200-299	0	332.42	0	290.41	0	290.20	0	287.12	0	571.20	0	559.46	0	802.52	0	744.38
	300-399	0	189.51	0	167.84	0	166.11	0	163.93	0	326.37	1	318.98	2	464.99	0	435.70
400+	1	294.93	1	258.27	0	258.18	0	256.69	0	503.39	0	493.44	0	713.01	0	668.89	
35-49	0	0	15024.99	1	13087.56	0	12842.95	4	12540.53	1	24242.45	2	23006.21	0	31259.50	5	26281.09
	1-9	0	8280.99	0	7220.45	1	7098.47	0	6949.45	0	13443.56	1	12702.30	1	17306.43	3	14707.59
	10-49	0	6411.51	0	5604.31	0	5503.22	0	5368.14	1	10365.28	0	9842.83	1	13469.39	1	11299.38
	50-99	0	1667.95	0	1453.64	0	1430.68	0	1399.73	0	2685.38	0	2505.71	0	3385.47	2	2838.65
	100-199	0	1122.99	0	973.32	0	956.37	1	940.63	0	1829.43	0	1727.57	1	2329.14	0	1922.39
	200-299	0	383.69	0	333.91	1	327.56	0	314.51	0	609.17	0	573.45	0	768.74	0	610.58
	300-399	1	189.79	0	162.04	0	158.09	1	153.55	0	290.75	1	256.95	0	331.57	0	262.22
400+	0	276.94	0	240.40	0	240.17	1	234.15	0	446.73	1	417.53	2	528.44	1	380.12	
50-60	0	0	6520.26	0	5507.26	1	5271.08	0	4953.74	0	8901.91	0	7436.69	0	8119.49	0	4368.22
	1-9	0	3660.71	0	3114.26	0	2977.16	0	2783.70	0	4991.06	0	4020.19	0	4313.84	0	2459.55
	10-49	0	2594.48	0	2220.12	0	2129.37	0	2023.49	0	3652.15	1	3024.64	0	3330.94	0	1913.74
	50-99	0	757.00	0	655.88	0	617.38	0	589.64	0	1079.67	0	840.55	0	882.80	0	480.45
	100-199	0	431.08	0	363.16	0	338.67	0	315.33	1	558.60	0	428.46	0	404.81	0	205.26
	200-299	0	150.84	0	129.23	0	124.18	0	113.13	0	198.93	1	184.87	0	183.18	0	85.64
	300-399	0	64.91	0	52.09	0	50.23	0	47.46	0	88.44	1	69.80	0	66.35	0	41.52
400+	0	85.66	0	73.02	0	65.30	0	55.61	0	99.95	0	85.21	0	86.28	0	40.53	

APPENDIX TABLE 1 Continued 付録 表 1 続き

Age at Exposure	Dose	Time Period															
		Oct. 1950- Dec. 1952		Jan. 1953- Dec. 1954		Jan. 1955- Dec. 1956		Jan. 1957- Dec. 1958		Jan. 1959- Dec. 1962		Jan. 1963- Dec. 1966		Jan. 1967- Dec. 1972		Jan. 1973- Dec. 1978	
		C	PY	C	PY	C	PY	C	PY	C	PY	C	PY	C	PY	C	PY
Nagasaki																	
0-9	0	0	1869.94	0	1648.68	0	1644.06	0	1638.07	1	3267.40	0	3249.22	0	4834.71	0	4804.85
	1-9	0	3686.34	0	3262.40	0	3255.92	0	3242.12	0	6457.24	1	6430.97	0	9592.43	1	9504.95
	10-49	0	2376.90	0	2103.70	0	2098.79	0	2091.94	0	4173.64	0	4160.17	0	6200.01	0	6132.52
	50-99	0	615.42	0	542.33	0	541.61	0	535.29	0	1064.78	0	1059.40	0	1583.56	0	1576.56
	100-199	0	390.05	1	344.39	0	344.23	0	343.77	0	684.78	0	683.20	0	1014.27	0	1013.07
	200-299	1	233.18	0	205.57	0	204.14	1	202.35	0	400.00	0	397.19	0	581.25	0	565.84
	300-399	0	90.18	0	79.94	0	80.06	0	79.94	0	160.00	0	160.00	0	240.06	0	239.78
	400+	0	166.84	1	146.55	0	139.70	1	137.61	0	272.00	0	272.00	0	408.10	0	407.62
10-19	0	0	2638.44	0	2308.62	0	2301.69	0	2284.03	0	4528.73	0	4493.53	0	6654.04	0	6542.53
	1-9	0	4396.23	0	3874.08	1	3852.16	1	3823.97	0	7589.02	0	7516.11	0	11092.23	0	10801.51
	10-49	0	1999.24	0	1762.34	0	1752.00	0	1741.08	0	3478.04	0	3461.55	0	5133.11	0	5044.37
	50-99	0	974.00	0	859.10	0	858.54	0	852.82	0	1702.79	0	1692.51	0	2504.83	0	2462.87
	100-199	0	1268.42	0	1119.51	0	1115.59	0	1108.26	0	2201.13	0	2183.35	0	3238.55	0	3174.54
	200-299	1	633.05	0	558.24	1	557.41	0	554.11	0	1098.52	1	1085.44	0	1602.33	0	1570.83
	300-399	0	177.24	0	154.93	0	154.11	0	153.89	0	305.87	0	304.00	0	456.06	0	431.96
	400+	2	234.22	0	205.85	0	206.15	0	204.81	0	404.37	0	402.95	0	592.82	0	574.45
20-34	0	0	1764.41	0	1550.88	0	1533.31	0	1518.15	0	3007.46	0	2954.00	0	4278.30	0	4060.22
	1-9	0	2830.49	0	2486.04	0	2463.06	0	2432.07	0	4797.33	0	4695.92	1	6900.38	1	6632.30
	10-49	0	1568.02	0	1373.90	0	1364.58	0	1349.79	0	2678.00	0	2643.44	0	3859.01	0	3715.18
	50-99	0	665.90	0	588.79	0	585.50	0	578.87	0	1148.14	0	1127.24	0	1641.74	0	1565.58
	100-199	1	668.01	0	583.20	0	580.17	0	575.60	0	1144.59	1	1121.42	0	1634.52	0	1539.39
	200-299	0	372.75	0	323.63	0	320.88	0	315.77	0	628.00	0	619.48	0	898.62	0	836.42
	300-399	0	160.07	0	141.11	0	140.09	0	139.91	0	280.00	0	279.62	0	403.41	0	395.28
	400+	1	171.84	0	149.90	0	147.12	0	141.80	0	280.00	0	277.12	0	402.23	0	387.89
35-49	0	0	1692.00	0	1458.72	0	1442.38	0	1404.46	0	2708.84	0	2575.14	0	3507.95	1	2849.04
	1-9	0	3036.81	0	2631.09	0	2585.92	0	2525.26	0	4811.03	0	4459.69	1	5855.87	0	4819.08
	10-49	0	1775.62	0	1534.31	0	1503.86	0	1465.02	0	2780.58	0	2605.94	0	3446.93	0	2872.64
	50-99	0	615.31	0	537.03	0	534.06	0	524.98	0	1016.87	0	951.26	0	1224.08	0	981.89
	100-199	0	517.52	0	452.15	0	440.81	0	429.21	0	837.84	0	770.71	0	1015.12	0	826.63
	200-299	0	262.86	0	227.53	0	224.15	0	221.04	0	426.64	1	398.02	0	530.36	0	429.36
	300-399	0	130.77	0	115.92	0	112.90	0	108.16	0	197.22	0	181.97	0	243.21	0	199.11
	400+	0	182.34	0	155.26	0	152.58	1	142.75	0	271.27	0	254.64	0	336.25	0	252.40
50-60	0	0	644.03	0	536.68	0	498.57	0	462.66	0	809.24	0	659.53	0	701.98	0	334.53
	1-9	0	1316.25	0	1105.06	0	1042.95	0	971.56	0	1742.04	0	1413.16	1	1465.60	0	752.95
	10-49	0	808.65	0	670.43	0	627.54	0	590.34	0	1030.01	0	810.56	0	822.72	0	401.71
	50-99	0	229.53	0	195.79	0	190.27	0	177.80	0	327.13	0	262.65	0	297.22	0	148.23
	100-199	0	190.23	0	155.35	0	149.39	1	138.52	0	239.62	0	192.02	0	220.15	0	113.95
	200-299	0	81.70	0	71.95	0	68.82	0	62.43	0	105.39	0	90.29	0	94.02	0	55.25
	300-399	0	40.58	0	33.13	0	30.55	0	28.02	0	54.12	1	44.97	0	35.68	0	21.43
	400+	0	49.85	0	41.97	0	40.93	0	37.97	0	68.39	0	49.21	0	42.31	0	20.30

REFERENCES

参考文献

1. BEEBE GW: Reflections on the work of the Atomic Bomb Casualty Commission in Japan. *Epidemiol Rev* 1:184-210, 1979
2. BEEBE GW, USAGAWA M: The major ABCC samples. ABCC TR 12-68
3. BEEBE GW, KATO H, LAND CE: Studies of the mortality of A-bomb survivors. 4. Mortality and radiation dose, 1950-66. *Radiat Res* 48:613-49, 1971 (ABCC TR 11-70)
4. MILTON RC, SHOHOJI T: Tentative 1965 radiation dose estimation for atomic bomb survivors, Hiroshima and Nagasaki. ABCC TR 1-68
5. KATO H, SCHULL WJ: Studies of the mortality of the A-bomb survivors. 7. Mortality, 1950-78: Part 1. Cancer mortality. *Radiat Res* 90:395-432, 1982 (RERF TR 12-80)
6. US DEPARTMENT OF ENERGY: Reevaluations of the dosimetric factors, Hiroshima and Nagasaki. DOE Symposium Series 55, DOE Technical Information Center, Washington DC, 1982.
7. KATO H, SCHULL WJ: Life Span Study Report 9. 1950-78. Supplementary Tables for RERF TR 12-80 and RERF TR 5-81
8. NATIONAL RESEARCH COUNCIL: Report of the Committee on the Biological Effects of Ionizing Radiation: The effects on populations of exposure to low levels of ionizing radiation (BEIR Report). National Academy of Sciences, Washington DC, 1980
9. NATIONAL COUNCIL ON RADIATION PROTECTION AND MEASUREMENTS: Influence of dose and its distribution in time on dose-response relationships for low-LET radiations: NCRP Report No. 64. Washington DC, 1980
10. ICHIMARU M, ISHIMARU T, MIKAMI M, YAMADA Y, OHKITA T: Incidence of leukemia in a fixed cohort of atomic bomb survivors and controls, Hiroshima and Nagasaki, October 1950-December 1978. RERF TR 13-81
11. ISHIMARU T, ICHIMARU M, MIKAMI M, YAMADA Y, TOKUNAGA Y: Distribution of onset of leukemia among atomic bomb survivors in the Leukemia Registry by dose, Hiroshima and Nagasaki, 1949-75. RERF TR 12-81
12. LAND CE, NORMAN JE: Latent periods of radiogenic cancers occurring among Japanese A-bomb survivors. *Late Biological Effects of Ionizing Radiation*, Vol. 1; Proceedings of a symposium, Vienna, 13-17 March 1978. Organized by the IAEA and WHO, Vienna, International Atomic Energy Agency, 1978. pp 29-74
13. LAND CE, TOKUNAGA M: Induction period. In *Recent Developments in Carcinogenesis: Epidemiological and Biological Significance*. Ed by Boice JD Jr., Fraumeni SF Jr., New York, Raven Press, 1983.
14. KALBFLEISCH J, PRENTICE RL: *The Statistical Analysis of Failure Time Data*. New York, Wiley, 1980 (Sections 3.5 and 5.2)
15. HOLFORD TR: The analysis of rates and survivorship using log-linear models. *Biometrics* 36:299-306, 1980
16. BAKER RJ, NELDER JA: *The GLIM system. Release 3. Generalized linear interactive modeling*. Numerical Algorithms Group, Oxford, 1978
17. COX DR: Regression models and life tables (with discussion). *J Royal Stat Soc B* 34:187-220, 1972

18. DOLL R: The age distribution of cancer: Implications for models of carcinogenesis. J Royal Stat Soc B 33:133-66, 1971
19. DOLL R, PETO R: The causes of cancer: Quantitative estimation of avoidable risks of cancer in the United States today. JNCI 66:1191-308, 1981
20. OTAKE M: Patterns of cancer mortality in the United States and Japan. RERF TR 13-79
21. HIRAYAMA T, Ed: Comparative Epidemiology of Cancer in the United States and Japan - Mortality. The United States-Japan Cooperative Cancer Research Program, Japan, 1977
22. EFRON B: Bootstrap methods: Another look at the jackknife. Ann Stat 7:1-26, 1977
23. EFRON B: The jackknife, the bootstrap, and other resampling plans. Stanford University, Technical Report No. 163, 1980