

Factors Affecting Recognition of Cancer Risks of Nuclear Workers

by

George W Kneale MA

Alice M Stewart FRCP

Department of Public Health
& Epidemiology
University of Birmingham
Edgbaston
Birmingham B15 2TT
UK

Tel: 0121 414 3367
Fax: 0121 414 3630

Sent to:
Occupational and Environmental
Medicine

January 1995

Abstract

Data from five nuclear facilities, whose dosimetry programmes were the responsibility of the US Nuclear Regulatory Commission, were included first in tests of what was needed to identify the cancer risks of badge monitored workers, and then in tests of whether strong central control of the US nuclear industry has achieved uniform standards of dosimetry.

From these tests, which were necessarily of much greater complexity than standard methods of cohort analysis, has come evidence that relations between exposure age and cancer risk are radically different for nuclear workers and A-bomb survivors, and evidence that the usual method of risk estimation - by linear extrapolation of high dose effects observed in A-bomb data - is both underestimating and distorting the cancer risks of nuclear workers. The tests have also shown that, even within nuclear facilities under NRC control, regional and temporal variations in dosimetry standards may be so great that pooling of data is not a safe option.

Introduction

Since 1977, when Mancuso and his associates first found evidence of a cancer risk for nuclear workers at Hanford,¹ there have been both confirmations and rebuttals of this occupational hazard.² On one side of this 'Hanford controversy' we have Kneale *et al*, whose findings are indicative of a risk which has left exposures after 50 years of age causing most of the extra cancer deaths after 70 years.³⁻⁵ On the opposite side we have Gilbert *et al* who have used standardised mortality ratio analyses to show that Hanford workers had low rates of cancer mortality, and cohort analyses to show that risk estimates based on A-bomb data are directly applicable to nuclear workers.⁶⁻⁹ According to A-bomb data the cancer risk from repeated exposure to small doses of radiation should be too small to show in Hanford data, and exposures after 50 years of age should be less dangerous than earlier exposures.¹⁰ Therefore, when recently confronted with "evidence of an increase in the excess relative risk with increasing age" Gilbert and her associates immediately suspected biased dosimetry.¹¹ They did, however, admit that "additional analyses addressing the modifying effects of factors such as age at exposure, time since exposure, calendar period of exposure, age at risk, birth cohort and calendar year of risk would be desirable".

The Hanford controversy is important since the only alternatives to A-data are occupational data, and biased dosimetry is important since the International Agency for Research on Cancer (IARC), is currently basing risk estimates for "carcinogenic effects of protracted low-dose exposures to radiation" on pooled data from USA, Canada and UK.¹² Therefore, included in the present report are data from five nuclear facilities, together with the results of including pooled and unpooled data in models of relative risk whose parameters included lag period, exposure age and exposure year.

Data

From the Hanford Environmental Health Foundation and the Oak Ridge Associated Universities were obtained the records of 85,642 badge monitored workers from five nuclear facilities: viz: Hanford, and four locations in or near Oak Ridge (i.e. X10, or the Oak Ridge National Laboratory, Y12, K25 and Fernald). These data were first divided into three series (Table I). In the first series were all the workers who had ever been at Hanford (cohort H, with 35,868 workers and 1,907 cancer cases). In the second were all the remaining workers who had ever been at X10 (cohort X, with 22,239 workers and 430 cancer cases), and in the third were all the residual workers (cohort Y, with 27,535 workers and 639 cancer cases). In addition to these cohorts, two sets of pooled data were obtained by first combining all the non-Hanford data (cohort XY with 49,774 workers and 1,069 cancer cases), and then adding the Hanford cohort (cohort HXY with 85,642 workers and 2,976 cancers).

For Hanford workers there were occupational data but no pay status records, and for other workers there were pay status records but no occupational data. Therefore, for the six levels of socio-economic status in Table II there are different criteria for Hanford and elsewhere. Several of the essential controlling factors in this Table were needed to cope with obvious differences between the five facilities. For example, at Hanford there were deaths to the end of 1986; elsewhere there were only deaths to the end of 1984, and births before 1900 were much commoner at Hanford (5%) than elsewhere (1.5%). The final year of dose recording was always the same (1978) but there were different starting dates, ranging from 1943 for X10 to 1954 for Y12. Until 1960 the average annual dose was higher for X10 than elsewhere, but thereafter the lead was taken by Hanford and, for the whole period, the average annual dose rate was higher for this cohort (2.9 mSv) than for X (1.55 mSv) or Y (0.70 mSv) (Table III). Finally, although

deaths from cancers of digestive organs were twice as common as deaths from genito-urinary (GU) cancers, among the 243 non-fatal cancers there were almost as many GU as lung cancers (Table IV).

Statistical Analysis

In many respects the method of statistical analysis was identical to the one recommended by Breslow and Day for cohort studies (Appendix).¹³ But over and above the procedures required for identification of suitable models of relative risk there were additional procedures to cope with the computer storage problems created by a need to observe the effects of adding extra parameters to a simple model of relative risk (Table V). As in the 1993 analysis of Hanford data⁵ the main parameter of each model was expressed as a doubling dose (β), and for the exponent of dose response (ϵ) the expected value (assuming linearity) was 1.0. One of three parameters for 'cancer modulating factors' was common to each model, namely, cancer latency or lag period (δ), and only the simplest of four models (model I) had no allowance for exposure age (α). The fifth parameter (which was also the fourth dose-weighting factor) was exposure year (γ). This factor was common to two of the three remaining models (III and IV), but only with model IV was there a full compliment of five parameters.

The first ever cohort analysis of Hanford data was the one by Kneale *et al* in 1981.³ At that time it was necessary for levels of monitoring for internal radiation to take the place of socio-economic status since neither pay status nor occupational data were available. In addition, there was smooth curve assessment of the cancer latency effect and the exposure age effect (fig. 1), but there was no allowance for what has since been suspected, namely, grossly different standards of dose recording in earlier and later years.¹⁴ In the 1993 analysis the controlling factors were exactly the same as in the present analysis (Table II) and, on each

occasion, there was step function assessment of three cancer modulating factors (Tables V and VI). With these assessments it was possible to have 'critical values', or Yes/No determinations, and thus obtain appropriate constraints for 'cancer effective doses' (see the window in fig. 2). By the gradual addition of extra parameters to a base model it was also possible to keep track of each effect, and thus be in a position to explore any problems created by non-uniform standards of dosimetry, after identifying suitable models (see chi-squares in Table VI) . For example, the best fitting model (IV) was used to obtain risk estimates for different cohorts (Tables VII to X) before testing for differences between pooled and unpooled data (Table XI). Finally, since the 1991 analysis of Hanford doses had left an impression of much better dosimetry standards after than before 1960,¹⁴ certain unexplained differences between the 1981 and 1993 analyses of Hanford data were also explored (Tables XII and XIII).

Risk Model Selection and Results

With the simplest of the four risk models (i.e. model I which made no allowance for possible effects of exposure age) there were no significant chi-squares in Table VI. With the remaining models (each with ten chi-squares) there were more significant results for the two cohorts containing Hanford workers (H and HXY) than for the other cohorts, and more for fatal cancers than for all cancers. Only model IV had significant chi-squares for all five sets of fatal cancers, and only the three cohorts with more than a thousand fatal cancers (H, XY and HXY) had critical values for exposure year which allowed each exposure year to contribute to the window doses for these cases.

Within the 'model IV windows' created by critical values of δ and α , the doubling doses for H and HXY were 6.5 and 8.2 mSv (fatal cancers) or 5.5 and 10.3 mSv (all cancers). For these 'cancer effective doses' the exposure age constraints

were 58 and 62 years, and the cancer latency or lag period constants were 14 and 17 years. For other cohorts with X10 workers (X and XY) the doubling doses were usually much lower (1.1 and 3.6 mSv for fatal cancers, or 2.7 and 13.0 mSv for all cancers) and there were also wider windows. Thus, the exposure age constraints were 40 or 48 years and the lag period constraints were 19 or 21 years.

The chi-squares in Table VI were actually "improvements to twice the log-likelihood relative to the null hypothesis of no radiation effect" (see Appendix). For the seven analyses which had, in relation to model IV, significant chi-squares, the number of cancers with measurable window doses (see the high risk cases in Table VII) ranged from 160 for HXY to 14 for Y. For cohort H the number of high risk cases was higher for all cancers (42) than for fatal cancers only (34), but for the fully pooled data (HXY) the number was lower for all cancers (119) than for fatal cancers (160). The proportion of high risk cases was higher for X10 (15.5%) than Hanford (2.3%) despite the fact that the average dose was higher for Hanford than X10 (Table III). For numbers of radiogenic cancers there were estimates which ranged from 84.9 for HXY to 10.8 for Y, and these too were much higher for X10 (29.2) than for Hanford (14.5).

The total number of high risk cases to feature in one or more of the seven model IV analyses was 346, with 190 from Hanford, 97 from X10 and 59 from elsewhere (see footnote to Table VII). It was eventually discovered that the two cohorts formed from pooled data (XY and HXY) were not on par with the three smaller cohorts (see below), so it was not possible to obtain corresponding numbers of radiogenic cancers. There was, however, no mistaking the fact that the cohort with the highest average dose (H) had a smaller proportion of radiogenic cancers (1.0%) than either X (9.6%) or Y (2.3%).

Further results of applying model IV to all fatal and non-fatal cancers are

shown in Table VIII, where observed and expected numbers of cancer cases (assuming no radiation effects) are given positions on an eight point scale of window doses. In addition to showing the effects of having a much less restricted age range for the window doses of X than H (see Table VI), this arrangement of the data shows that both for dose-weighted and rank-weighted \underline{t} values, there were significant values even for the smallest cohort (which also had the lowest average dose). These findings were clearly the result of there being a dose-related cancer risk even when the average dose was less than 1 mSv per annum (Table III). Thus, for window doses equal to or greater than 3 mSv, the ratio of observed to expected cancer cases was actually higher for the cohort with the lowest average dose (Y cohort, with 4 cases observed and 1.9 expected) than for Hanford (25 observed and 15.5 expected) or X10 (48 observed and 39.5 expected).

Given the exposure age constraints of model IV (and the relatively large numbers of Hanford workers who were born before 1900) it was inevitable that the high risk cases would be older than average, and that this bias would affect Hanford more than elsewhere (Table IX). Also shown in this table is the high proportion of non-fatal cancers for high risk cases and cancer deaths after 70 years of age, and the fact that the cohort with the longest follow-up period (Hanford) accounted for 84% of the cancer deaths after 80 years. This cohort included 72% of the non-fatal cancers, 70% of the births before 1900, and 100% of the deaths after 1984. But instead of the proportion of high risk cases being higher for Hanford than elsewhere, the reverse was true (i.e. Hanford 2.3% and elsewhere 8.0%).

Finally, at Hanford and elsewhere there was a preponderance of GU cancers both among the high risk cases and among the non-fatal cancers (Table X). These biases were probably the result of prostate tumours having an exceptionally good

prognosis, since in other (histological) respects there was remarkably little difference between the high risk cases and the other cancers.

Effects of Pooling Data from Different Facilities

The method of identifying cancer deaths (originally devised by Mancuso)¹, was exactly the same for each cohort. In theory, all facilities controlled by the US Nuclear Regulatory Commission have the same radiation monitoring programmes and methods of dose estimation. But earlier work had shown that, in these respects, there was considerable scope for differences between different locations and decades.^{14,15}

With comparable standards of dose-recording in each facility and each calendar year, the risk estimates would be the same for pooled and unpooled data. Furthermore, with any good fitting model, the sum of two separate log-likelihoods would be equal to the single log-likelihood for the combined data. Therefore, in Table XI are shown, for each model, a) the chi-squares for two sets of pooled data (cohort XY and HXY), b) the chi-squares for two components of these sets (either X and Y, or H and XY), c) the differences between the chi-squares for matching sets of pooled and unpooled data, and d) the critical chi-square values needed to establish significant difference between the matching sets.

For model I none of the chi-square differences for pooled and unpooled data came anywhere near the critical difference. For the smaller of the two data sets (X + Y and XY) only model IV succeeded in establishing any significant differences between the pooled and unpooled data, but for the larger set, (H + XY and HXY) there was definite evidence of a difference with models II and III, and suggestive evidence with model IV.

Differences Between the 1981 and 1993 Analyses of Hanford Data

Certain unexplained differences between the 1981 and 1993 analyses of Hanford data were the original reason for introducing two of the model IV parameters (ϵ and γ). These parameters were needed since, in the earlier analysis, definite evidence of a dose-related effect for a large group of (A) cancers was accompanied by negative dose trend for the remaining (B) cancers. Therefore, for all cancers there was no certain evidence of any cancer effect. Furthermore, for A cancers, there was definite evidence of non-linearity of dose response, with an exponent of dose response well below unity. Neither of these findings was confirmed in the 1993 analysis. Therefore, bearing in mind the possibility of much better recording of radiation doses after than before 1960,¹⁴ a systematic search of the whole parameter space was made to discover whether, in addition to a global maximum of log-likelihood, there was also a local maximum that had more in common with the 1981 than the 1993 analysis.

Before this search was made it was arguable that controlling for levels of internal monitoring on the first occasion, and controlling for socio-economic status on the second occasion, was sufficient to account for the different findings of the 1981 and 1993 analyses. However, the search both revealed a local maximum and showed that, even with the same controlling factors, there would have been significant differences between the 1981 and 1993 analyses (Table XII). Thus, with the local maximum, the critical values were 50 years (for exposure age) and 1956 (for exposure year) and with the global maximum the corresponding values were 62 years and 1979. Furthermore, the ϵ values for the exponent of dose response were much lower with the local maximum (0.17 or 0.02) than with the global maximum (1.31 or 1.14).

If the correct value of ϵ were as low as 0.2 there would be both a sizable cancer risk at the lowest dose level and little change with increasing dose. For example, with a window dose of 0.1 mSv, the relative risk would be 1.37 for fatal cancers and 1.87 for all cancers, and a thousand fold increase (to 100 mSv) would only increase the relative risk to 2.05 or 2.22. Therefore, the absurdly low values of ϵ with the local maximum were probably the result of there being a time when failure to record more than a fraction of the true doses of plumbers and process workers at Hanford was producing falsely small differences between the highest and lowest annual doses. In line with this suggestion is the 1991 analysis of Hanford doses¹⁴ which showed that, before 1960, there were several years when the average annual dose was barely a tenth of later averages.

Discussion

The results of the present analysis are difficult to reconcile either with the assumption that the cancer experiences of A-bomb survivors are a reliable source of risk estimates for nuclear workers, or with the assumption that the pooled data of IARC will prove to be a satisfactory alternative to A-bomb data. The A-bomb data are unsatisfactory because relations between exposure age and cancer risk are manifestly different for survivors and workers, and the pooled data of IARC are unsatisfactory because even within one source of these data (USA) there is evidence of cohort heterogeneity.

A slow unfolding of the mortality experiences of A-bomb survivors has repeatedly left statisticians with an impression of a) no late effects of the bombing apart from a few extra cancer deaths; b) no cancer risk below a certain dose level, and c) a smaller cancer risk for persons who were over 50 years when exposed than for younger survivors. Therefore, observers of the Hanford controversy, who

included the US Committee on Biological Effects of Ionizing Radiation (BEIR), have repeatedly sided with the rebuttals. However, if the experiences of A-bomb survivors were a true guide to the cancer risks of nuclear workers, the effects of including exposure age among the parameter of a relative risk model would have been very different from the observed effects. Likewise, if there had been uniform standards of dosimetry in all nuclear facilities under NRC control, neither the comparisons between pooled and unpooled data in Table XI nor the comparisons between the global and local maxima in later tables, would have revealed any significant differences.

Some idea of the extent to which the model IV risk estimates differ from estimates based on A-bomb survivors can be gleaned by comparing the BEIR V and model IV estimates.¹⁶ According to BEIR V, if 100,000 persons with an average life span of 65 years had a continuous lifetime dose of 1.0 mGy/y they would probably experience 990 extra cancer deaths. With no such exposures the expected number of cancer deaths would be 20,100. Therefore, on this basis the average doubling dose would be close to $(20100/990) \times 65$ or 1320 mGy. For the largest of five cohorts (HXY) model IV had 8.2 mSv as the doubling dose for fatal cancers, 58 years as the critical exposure age, and 14 years as the critical lag period (Table VI and fig. 1). Therefore, by 66 years of age an annual dose of 1.0 mSv would be equivalent to a window dose of 8.0 mSv, and the number of cancer deaths after 80 years would be twice the expected number. Nowadays, both in the United States and in Britain, cancer deaths after 80 years account for a quarter of all fatal cancers. Therefore, according to model IV, the average doubling dose would be close to 8.2×4 or 33 mSv.

These are necessarily rough comparisons, but they are a reminder that it is only after 50 years of age that the model IV estimates are much higher than the

BEIR V estimates (fig. 1). They are also a reminder that although Kneale *et al* are not alone in producing higher estimates of relative risk for nuclear workers than for A-bomb survivors - similar findings have been reported by Wing *et al* in the US¹⁷ and by Kendal⁰ *et al* in the UK,¹⁸ - no one else has used a risk model which allows for cancer modifying effects of exposure age and exposure year. Failure to make any allowance for exposure age effects is clearly the reason why both the Gilbert *et al* analysis of Hanford data, and the IARC analysis of data from USA, Canada and Britain, failed to find any evidence of extra radiogenic cancers. For recognition of the necessarily small cancer risks from the strictly controlled doses of nuclear workers it may also be important to know where and when the doses were recorded, since the model IV analysis has shown that, even with strong central control of the US nuclear industry, it was not possible to maintain uniform standards of dosimetry either in different facilities or in the same facility at different points in time.

The strain placed on any cohort analysis by a need to consider several 'cancer modulating factors' is immense. Nevertheless, by making suitable additions to routine procedures, it has been possible to obtain a much clearer impression of what is happening in nuclear facilities, and what may be happening elsewhere as a result of there being both continuous and universal exposure to natural background radiation.

Appendix

Up to the point of formal derivation of the optimal test of the four relative risk models, the statistical procedures were identical with the standard (Breslow and Day) methodology.¹³ But when it came to practical application of these results for computation of plausible models with the currently available data, there were many unsolved problems. For example, with all the possible confounders in Table II the number of risk sets ran into millions. It was also possible for the same individual to reappear in several risk sets, and this redundancy made the number of controls in each risk set so large that direct methods of calculation were impossible. Furthermore, one alternative to the null hypothesis (of no cancer effects of the radiation exposures) was that the risk increases linearly with dose. Therefore the risk model was more complex in computation than the standard logistic model (especially when it came to the calculation of differential coefficients). Finally, one of the possible models for cancer latency effects was that all doses within a critical pre-lag period had no effect. This would make differential coefficients for variation of likelihood with critical interval formally non-existent (since lag was necessarily measured in whole years). So there was a need to consider methods of calculating the maximum of the likelihood function that did not depend upon estimation of differential coefficients. Standard statistical packages such as PECAN could not easily be modified to resolve these problems since the situation was one which required separate identification of the effects of four dose-weighting factors. Hence the need for the following additions to a standard cohort analysis.

Computational Difficulties:

A) Number of Risk Sets

The nuisance parameters, together with their ranges, are given in Table II. Multiplying these ranges together (to obtain the total number of possible risk sets)

produced a figure of 4,293,120. This was much too large for computer storage. However, only the risk sets with cancer deaths were actually 'informative' ones, and, in a table of all possible risk sets, the informative sets were so sparse that they could be indexed by the hash technique of Knuth.¹⁹ This reduced the storage requirement to slightly more than twice the number of cancer cases (i.e. less than 6,000).

B) Large Risk Sets

The number of selections of n_s objects taken without replacement from N_s objects is of the order of $N_s^{n_s}$, can be very large even when N_s and n_s are moderate. This number determined the number of terms in the denominator sum of products for the contribution to the likelihood (L) of the risk sets (s), and sets containing more than 5 cancer cases and more than 100 cases and matched controls were quite common. Therefore, direct calculation of L by the Breslow and Day formula (see page 186 of Vol II)¹³ was often impossible. Furthermore, direct calculation would have required storage of too many calculated values of relative risk. However, an alternative was found by realising that, since a symmetric function of relative risk was involved, these risks could be calculated from power sums in much the same way as k-statistics can be calculated from moments.²⁰ A suitable formula was derived as follows:

Let: i index the individuals in a risk set with n cases, with a total of N cases and matched controls;

R_i be the calculated relative risk of individual i ;

S_p be the sum of the powers (p) of the R_i , i.e. $S_p = \sum_{i=1}^N R_i^p$, and

D_p be the sum of the products over all selections of p taken without replacement from the total of N , and only ordered selections taken

into account.

$$\text{Thus: } D_1 = \sum_i R_i, D_2 = \sum_{i < j} R_i R_j, D_3 = \sum_{i < j < k} R_i R_j R_k \quad \text{etc.}$$

Then by inspection $D_1 = S_1$, $2D_2 = S_1^2 - S_2$ etc. and the following recursive result can be verified by the principle of alternating exclusion and inclusion:

$$pD_p = \sum_{q=1}^p (-1)^{(q+1)} D_{p-q} S_q \quad \text{where } D_0 \text{ is defined as unity. This left } D_n \text{ as the required denominator.}$$

Also important is the fact that this result only required storage of n_s power sums S_q for each risk set, where n_s was the maximum number of cases in the risk set. This meant that, together with the efficient storage of risk sets provided by the hash table technique, the total computer storage was not excessive.

C) Non-differential Likelihoods

In the formula for the risk set contribution to L the derivation of the denominator from power sums of relative risks for risk sets made it hard to find the differential coefficients of L with respect to the parameters (α , β , γ , δ and ϵ) even though L itself is easily calculated. This meant that variants of the Newton-Raphson algorithm could not be used to find the values of the parameters that maximises L . Instead, L was maximised by varying the parameters directly and a suitable algorithm was the simplex one of Nelder and Mead.²¹

Method of the Resulting Computer Program

A first pass through the data initialised the hash table for storage of the power sums of the informative risk sets. Then $-2\ln(L)$ was calculated for each successive approximation to the minimising parameters, by a single pass through the data. For each member of each study cohort the following procedures were used:

If he or she was a case then $-2\ln(R)$ was added to $-2\ln(L)$, where R was the relative risk in the death year. Since the same individual might be a control in other informative risk sets; each employment year of the hash table was scanned to see which risk sets with potential death years were informative and, where necessary, appropriate additions were made to the power sums. At the end of these passes through the data a sequential pass was made through the hash table (in order to calculate the denominator of each informative risk set from corresponding power sums) before updating the value of $-2\ln(L)$ by any contribution from the denominators. Finally, as a minor improvement, a constant for each risk set (depending on n_s and N_s) was added to $-2\ln(L)$ so that the contribution to $-2\ln(L)$ was zero when all the relative risks in a given risk set were equal to unity. In other words, the final value of $-2\ln(L)$ was the approximate chi-square referred to in the text.

References

1. Mancuso TF, Stewart AM, Kneale GW. Hanford I: Radiation exposures of Hanford workers dying from cancer and other causes. *Health Physics* 1977;**33**:369-384.
2. Stewart AM, Kneale GW. An overview of the Hanford controversy. Occupational Medicine: State of the Art Reviews. Philadelphia: Hanley and Belfus Inc 1991;**6**:641-663.
3. Kneale GW, Mancuso TF, Stewart AM. Hanford radiation study III: A cohort study of the cancer risks from radiation to workers at Hanford (1944-1977 deaths) by the method of regression models in life-tables. *Br J Ind Med* 1981;**16**:156-166.
4. Stewart AM, Kneale GW. The Hanford data: Issues of age at exposure and dose recording. The PSR Quarterly, Sept. 1993;**3**:101-111.
5. Kneale GW, Stewart AM. Reanalysis of Hanford data: 1944-86 deaths. *Am J Ind Med* 1993;**23**:371-389.
6. Gilbert ES, Marks S. An analysis of the mortality of workers in a nuclear facility. *Radiat Res* 1979;**79**:122-148
7. Gilbert ES, Petersen GR, Buchanan A. Mortality of workers at the Hanford site: 1945-1981. *Health Physics* 1989;**56**:11-25.
8. Gilbert ES, Fry S, Wiggs L, Voelz GL, Cragle DL, Petersen GR. Analyses of combined mortality data on workers at the Hanford Site, Oak Ridge National Laboratory, and Rocky Flats Nuclear Weapons Plant. *Radiat Res* 1989;**120**:1-17.
9. Gilbert ES, Omohundro E, Buchanan JA, Holter NA. Mortality of workers at the Hanford site: 1945-1986. *Health Physics* 1991;**64**:577-590.
10. Kato H, Schull WJ. Studies of the mortality of A-bomb survivors:
7. Mortality, 1950-1978: Part 1. Cancer Mortality. *Radiat Res* 1982;**90**:395-432.

- 11 Gilbert ES, Cragle DL, Wiggs LD. Updated analyses of combined mortality data for workers at the Hanford Site, Oak Ridge National Laboratory, and Rocky Flats Weapons Plant. *Radiat Res* 1993;**136**:408-421.
- 12 IARC Study Group on Cancer Risk Among Nuclear Industry Workers. Direct estimates of cancer mortality due to low doses of ionizing radiation: an international study. *Lancet* 1994;**344**:1039-1043.
- 13 Breslow NE, Day NE. Statistical methods in cancer research: Vol. I. The analysis of case-control studies. Vol. II. The design and analysis of cohort studies. Lyon: 1980 IARC Sci Publi No. 32.
- 14 Kneale GW, Sorahan TM, Stewart AM. Evidence of biased recording of radiation doses of Hanford workers. *Am J Indust Med* 1991;**20**:799-803.
- 15 Wing S, Shy CM, Wood JL, Wolf S, Cragle DL, Tankersley W, *et al.* Job factors, radiation and cancer mortality at Oak Ridge National Laboratory: Follow-up through 1984. *Am J Ind Med* 1993;**23**:265-279.
- 16 BEIR V *Health effects of exposure to low levels of ionizing radiation.* 1990 National Academy Press, Washington DC.
- 17 Wing, S, Shy CM, Wood JL, Wolf S, Cragle DL, Frome EL. Mortality among workers at Oak Ridge National Laboratory: Evidence of radiation effects in follow-up through 1984. *JAMA* 1991;**265**:1397-1402.
- 18 Kendall GM, Muirhead CR, MacGibbon BH, O'Hagan JA, Conquest AJ, Goodill AA, *et al.* Mortality and occupational exposure to radiation: First analysis of the National Registry for Radiation Workers. *BMJ* 1992;**302**:220-225.
- 19 Knuth DE. The art of computer programming, sorting and searching, Vol 3. London: Addison-Wesley Publication Co 1975.

- 20 Kendall MG, Stuart A. The advanced theory of statistics, Vol 1.
Distribution Theory. London: Charles Griffin & Company Ltd 1963.
- 21 Nelder JA, Mead R. A simplex method for function minimization.
Comput J 1965;**7**:308-313

Table I Specifications of five study cohorts of nuclear workers

Cohort	Principal Work Place	Workers ¹	Cancer Cases		Total
			Fatal	Non-fatal ²	
H	Hanford	35,868 (2772)	1732	175	1907
X	X10*	22,239 (2911)	401	29	430
Y	Y12	14,611	277	17	294
	K25	7,524 (2823)	180	17	197
	Fernald	5,400	143	5	148
XY	Oak Ridge Fernald	49,774 (5734)	1001	68	1069
HXY	Hanford Oak Ridge Fernald	85,642 (8506)	2733	243	2976

1 () Workers in more than one facility

2 Non-fatal cancers with other stated causes of death

Starting dates: Hanford 1944; X10 1943; K25 1945; Fernald 1952; Y12 1954.

Final dates: (1) for death ascertainment: Hanford 1986 Elsewhere 1984
(2) for recorded doses: 1978 all workers

* probably synonymous with Oak Ridge National Laboratory (ORNL)

Table II Essential controlling factors for the cohort analysis

Factor	Levels	Details	
Sex	2	Male ; Female	
Race	2	White ; other	
Birth Year	20	5 year intervals: 1870 to 1964	
Hire Year	13	2 year intervals: 1944 to 1978	
Facility	17	(1) X10 only (2) Y12 (3) X10 + Y12 (4) K25 only (5) K25 + X10 (6) K25 + Y12 (7) K25 + Y12 + X10 (8) Fernald only (9) Fernald + X10 (10) Fernald + Y12 (11) Fernald + Y12 + X10 (12) Fernald + K25 (13) Fernald + K25 + X10 (14) Fernald + K25 + Y12 (15) Fernald + K25 + Y12 + X10 (16) Hanford only (17) Hanford and elsewhere	
Potential Year of Death ⁽¹⁾	43	1 year intervals: 1944 to 1986	
Socio-Economic Status ⁽¹⁾	6	H Cohort	Other Cohorts
		(1) Professional	Monthly salary
		(2) Managerial	Contract salary
		(3) Clerical	Other salary
		(4) Craftsmen	Weekly wage
		(5) Other blue collar	Hourly wage
		(6) Not specified	No record
Discharge Status ⁽¹⁾	17	Died at work: yes/no If still alive: Still working after age 60 years yes/no Discharge status recorded yes/no Employment more than 3 years yes/no Post employment period 3+ years yes/no	

(1) separate assessment for each calendar year

Table III Three distinct cohorts. Average workforce (W) and average annual radiation doses (R) for seven consecutive periods

Period	Cohort H		Cohort X		Cohort Y	
	W	R	W	R	W	R
		mSv		mSv		mSV
pre 1950	5379	1.02	2116	1.87	182	1.19
1950-54	7773	1.50	3635	2.13	1150	0.95
1955-59	8769	2.40	4464	3.53	2720	1.81
1960-64	8361	4.47	4975	1.94	5597	1.01
1965-69	8379	4.86	5875	1.20	6166	0.82
1970-74	7127	3.28	5546	0.69	6629	0.66
1975-78	9146	2.37	7319	0.39	10926	0.31
1943-78	7740	2.92	4620	1.55	5696	0.70

Table IV Primary sites of three series of fatal and non-fatal cancers

ICD Nos. 8th Revision	Cohort H		Cohort X		Cohort Y		Total	
	F	NF	F	NF	F	NF	F	NF
140-149 Mouth & pharynx	41	3	12	-	7	3	60	6
150-159 Gastro-intestinal	461	41	98	6	142	4	701	51
160-163 Respiratory	532	46	109	7	225	10	866	63
170-174 Bone & connective tissue	106	3	26	-	23	-	155	3
180-189 Genito-urinary	223	49	59	6	72	6	354	61
190-194 Brain & endocrine	59	4	11	2	27	3	97	9
195-199 Non-specific	128	11	32	2	36	2	196	15
200-203 Lymphomas	112	7	30	4	44	6	186	17
204-209 Leukaemias	70	11	24	2	24	5	118	18
Totals	1732	175	401	29	600	39	2733	243

F = Fatal cases

NF = Non-fatal cases

Table V Parameters of four models of relative risk

Model	Main Parameter		Extra Parameters		
	β	ϵ	δ^1	α^1	γ^1
I	E	E	E	DV	DV
II	E	E	E	E	DV
III	E	DV	E	E	E
IV	E	E	E	E	E

E = an estimated value by maximum likelihood

DV = default value

β = doubling dose in mSv

ϵ = exponent of dose-response

δ = minimum cancer latency or lag period

α = minimum age exposure

γ = latest exposure year

¹ = critical values marking the boundaries of
'risk set windows' (see fig. 2 and table VI)

Table VI Results of including five cohorts in four risk models

1. Fatal Cancers

Risk Model	Cohort	Estimated Parameters					Chi-square Tests	
		β	ϵ	Critical Values			χ^2 (df)	Significance
I	H	284	1.84	24+	-	-	0.87	ns
	X	<1	0.01	22+	-	-	5.06	ns
	Y	<1	0.01	20+	-	-	1.43 (3)	ns
	XY	<1	0.01	20+	-	-	6.05	ns
	HXY	5235	0.03	10+	-	-	1.27	ns
II	H	6.6	1.27	17+	62+	-	13.56	*
	X	5.2	0.46	18+	48+	-	8.08	ns
	Y	2.7	0.43	19+	48+	-	7.90 (4)	ns
	XY	3.6	0.39	19+	48+	-	13.71	*
	HXY	8.0	0.37	14+	58+	-	16.32	*
III	H	4.9	-	17+	62+	<1979	13.13	*
	X	15.7	-	18+	48+	<1962	6.25	ns
	Y	1.6	-	13+	54+	<1959	10.84 (4)	*
	XY	15.7	-	18+	48+	<1962	9.74	*
	HXY	2.4	-	25+	57+	<1959	8.11	ns
IV	H	6.5	1.31	17+	62+	<1979	13.56	*
	X	1.1	0.47	21+	45+	<1957	13.52	*
	Y	0.1	0.38	13+	54+	<1959	13.26 (5)	*
	XY	3.6	0.40	19+	48+	<1981	13.71	*
	HXY	8.2	0.37	14+	58+	<1981	16.32	*

2. Fatal and non-fatal cancers

I	H	323	2.04	24+	-	-	0.44	ns
	X	<1	0.01	22+	-	-	4.44	ns
	Y	16	0.01	20+	-	-	0.26 (3)	ns
	XY	4830	0.11	20+	-	-	3.41	ns
	HXY	307	0.001	10+	-	-	0.87	ns
II	H	5.5	1.14	17+	62+	-	14.42	*
	X	4.7	0.53	21+	45+	-	9.45	ns
	Y	6.6	0.42	22+	48+	-	3.12 (4)	ns
	XY	13.4	0.55	19+	48+	-	8.71	ns
	HXY	10.3	0.44	16+	59+	-	12.19	*
III	H	4.8	-	17+	62+	<1979	14.23	*
	X	19.3	-	18+	48+	<1963	6.48	ns
	Y	2.5	-	21+	54+	<1964	7.61 (4)	ns
	XY	22.5	-	18+	48+	<1962	7.90	ns
	HXY	2.9	-	25+	57+	<1959	7.71	ns
IV	H	5.5	1.14	17+	62+	<1979	14.42	*
	X	2.7	0.25	22+	40+	<1969	10.11	ns
	Y	3.6	1.32	21+	54+	<1970	7.87 (5)	ns
	XY	13.0	0.55	19+	48+	<1976	8.71	ns
	HXY	10.3	0.45	16+	59+	<1984	12.19	*

* = $p > 0.05$

ns = not significant

Table VII Model IV. Estimated numbers of high risk cases and radiogenic cancers

Series	Cohort ¹	Cancer Cases				
		Total ²	Informative ³	High Risk ⁴	Radiogenic ⁵	
		No.	No.	No.	No.	%
Fatal Cancers	H	1732	1476.7	34	14.5	1.0
	X	401	303.6	47	29.2	9.6
	Y	600	461.4	14	10.8	2.3
	XY	1001	765.0	88	43.3	5.7
	HXY	2733	2241.7	160	84.9	3.8
All Cancers	H	1907	1618.5	42	18.1	1.1
	HXY	2976	2436.8	119	57.1	2.3

1 = only cohorts with significant chi-squares for model IV are included (see table VI)

2 = see table I

3 = cancer cases remaining after risk set matching

4 = cases whose window doses equalled or exceeded the lowest recorded dose of 0.1 mSv. The total number of cases from the seven analyses was 346 (with 190 from cohort H, 97 from cohort X, and 59 from cohort Y) see tables IX and X

5 = estimates of the number of extra, radiogenic cancers and the proportion of informative cases

Table VIII Model IV. Observed and expected numbers of fatal and non-fatal cancers for eight window doses and five cohorts

Window Doses mSv ^a	Cohort H			Cohort X			Cohort Y			Cohort XY			Cohort HXY		
	Obs	Exp	t value	Obs	Exp	t value	Obs	Exp	t value	Obs	Exp	t value	Obs	Exp	t value
0.0	1865	1865.8	-0.22	347	358.2	-2.78	630	632.0	-1.16	947	984.3	-2.09	2927	2928.1	-0.27
0.1	1	2.56	-1.08	3	2.89	+0.09	0	0.22	-0.49	1	2.81	-1.17	2	3.14	-0.72
0.3	8	11.79	-1.40	12	11.66	+0.12	2	2.72	-0.64	17	15.89	+0.34	11	13.66	-0.91
1.0	8	11.33	-1.34	20	17.82	+0.70	3	2.14	+0.75	33	30.18	+0.75	10	14.58	-1.64
3.0	13	10.84	+0.97	24	21.23	+0.81	1	1.31	-0.32	23	20.34	+0.76	14	11.83	+0.92
10.0	10	4.03	+3.91	18	12.58	+1.92	2	0.42	+2.62	12	10.59	+0.53	10	4.09	+3.83
30.0	2	0.60	+1.86	2	4.30	-1.26	1	0.15	+2.29	8	4.35	+2.00	2	0.60	+1.86
100.0	0	0.00	0.00	4	1.34	+2.67	0	0.00	0.00	1	0.56	+0.64	0	0.00	0.00
Total	1907			430			639			1069			2976		
Dose Weighted t value	3.94 ***			2.40 **			3.21 ***			2.20 **			3.86 ***		
Rank-Weighted t value	2.31 **			3.15 ***			2.17 **			2.85 ***			2.05 **		

p < 0.01
*** p < 0.001

a = these are cancer effective doses. For the boundaries of the window doses see model IV (fatal and non-fatal cancers) in Table VI.

Table X Diagnostic categories of the model IV high risk cases

Facilities	ICD Nos.	Fatal and Non-Fatal Cancers			Fatal Cases Only		
		All Cases	High Risk Cases		All Cases	High Risk Cases	
		No.	No.	%	No.	No.	%
Hanford	140-149	44	2	4.5	41	1	2.4
	150-159	502	45	9.0	461	39	8.5
	160-163	578	61	10.6	532	52	9.8
	170-174	109	6	5.5	106	6	5.7
	180-189	272	42	15.4	223	33	14.8
	190-199	202	17	8.4	187	14	7.5
	200-209	200	17	8.5	182	14	7.7
	Total	1907	190	10.0	1732	160	9.2
Elsewhere	140-149	22	2	9.1	19	2	10.5
	150-159	250	46	18.4	240	42	17.5
	160-163	351	39	11.1	334	36	10.8
	170-174	49	3	6.1	49	3	6.1
	180-189	143	34	23.8	131	29	22.1
	190-199	115	17	14.8	106	17	16.0
	200-209	139	15	10.8	122	13	10.7
	Total	1069	156	14.6	1001	142	14.2

Table IX Model IV. Age distributions of fatal and non-fatal cancers.
Main series and high risk cases

Death Age Years	Hanford		Elsewhere	
	All Cancers	High Risk Cases ¹	All Cancers	High Risk Cases ¹
under 60	600 (26)	-	470 (9)	-
60-64	324 (13)	-	197 (17)	3 (1)
65-69	355 (29)	-	184 (14)	38 (1)
70-74	289 (36)	29	129 (15)	57 (7)
75-79	195 (38)	84 (11)	61 (8)	38 (3)
80-84	106 (22)	52 (12)	25 (3)	19 (2)
85+	38 (11)	25 (7)	3 (2)	1
Total	1907 (175)	190 (30)	1069 (68)	156 (14)
7	No. 628 (107)	190 (30)	218 (28)	115 (12)
	% 32.9 (61.1)	100.0 (100.0)	20.4 (41.2)	73.7 (85.7)

() = non-fatal cancers

¹ = see footnote to table VII

Table XI Tests of differences between two sets of pooled data (XY and HXY) and two components of each pool

1) Non-Hanford (X, Y and XY)

Series	Model	Cohorts			Differences Between XY and X+Y		
		XY Chi-square	X Values ¹	+ Y	Chi-Square Difference	Critical Difference	significant non-homogeneity
Fatal Cancers	I	6.05	5.06	+ 1.43	0.44 (3)	7.82	ns
	II	13.71	8.08	+ 7.90	2.27 (4)	9.49	ns
	III	9.74	6.25	+ 10.84	7.35 (4)	9.49	ns
	IV	13.71	13.52	+ 13.26	13.07 (5)	11.07	*
All Cancers	I	3.39	4.44	+ 0.26	1.34 (3)	7.82	ns
	II	8.71	9.45	+ 3.12	3.86 (4)	9.49	ns
	III	7.90	6.48	+ 7.61	6.19 (4)	9.49	ns
	IV	8.71	10.11	+ 7.87	9.27 (5)	11.07	ns

2) Hanford and elsewhere (H, XY and XXY)

Differences Between HXY and H + XY

		HXY			H + XY		
		HXY	H	+ XY			
Fatal Cancers	I	1.27	0.87	+ 6.05	5.65 (3)	7.82	ns
	II	16.32	13.56	+ 13.71	10.95 (4)	9.49	*
	III	8.11	13.13	+ 9.74	14.76 (4)	9.49	*
	IV	16.32	13.56	+ 13.71	10.95 (5)	11.07	ns
All Cancers	I	0.87	0.44	+ 3.41	2.94 (3)	7.82	ns
	II	12.19	14.42	+ 8.71	10.84 (4)	9.49	*
	III	7.71	14.23	+ 7.90	14.42 (4)	9.49	*
	IV	12.19	14.42	+ 8.71	10.84 (5)	11.07	ns

1 = see table VI
* = p < 0.05
ns = not significant

Table XII Cohort H. Comparisons between global and local maximums for the log-likelihood function

1) Model IV parameters

Maximum of Log Likelihood Function	Cases	Estimated Parametner					Estimated Numbers		Chi-Square values (5df)
		β	ϵ	δ	α	γ	EDC cases	Radiogenic cases	
Global	Fatal Cancers	6.5	1.31	17+	62+	<1979	34	14.3	13.56
	All Cancers	5.5	1.14	17+	62+	<1979	42	18.1	14.42
Local	Fatal Cancers	34	0.17	8+	50+	<1956	255	89.4	12.73
	All Cancers	93	0.02	8+	50+	<1956	317	154.2	14.91

2) Model IV dose trends

	Window Dose		Global Maximum				Local Maximum			
	mSv	Obs.	Exp.	\pm value	Obs.	Exp.	\pm value	Obs.	Exp.	\pm value
t value	0.0	1865	1865.8	-0.22	1580	1610.2	-3.90			
	0.1	1	2.56	-1.08	9	6.69	+0.97			
	0.3	8	11.79	-1.40	71	62.07	+1.41			
	1.0	8	11.33	-1.34	87	84.08	+0.41			
	3.0	13	10.84	+0.97	107	97.88	+2.14			
	10.	10	4.03	+3.91	34	30.63	+0.75			
	30.	2	0.60	+1.86	18	15.03	+0.89			
	100.	0	0.00	0.00	1	0.47	+0.81			
Dose-weighted		3.94 ***				1.78 ns				
Rank-weighted		2.31 **				3.46 ***				

Table XIII High risk cases in cohort H. Comparisons between global and local maximums

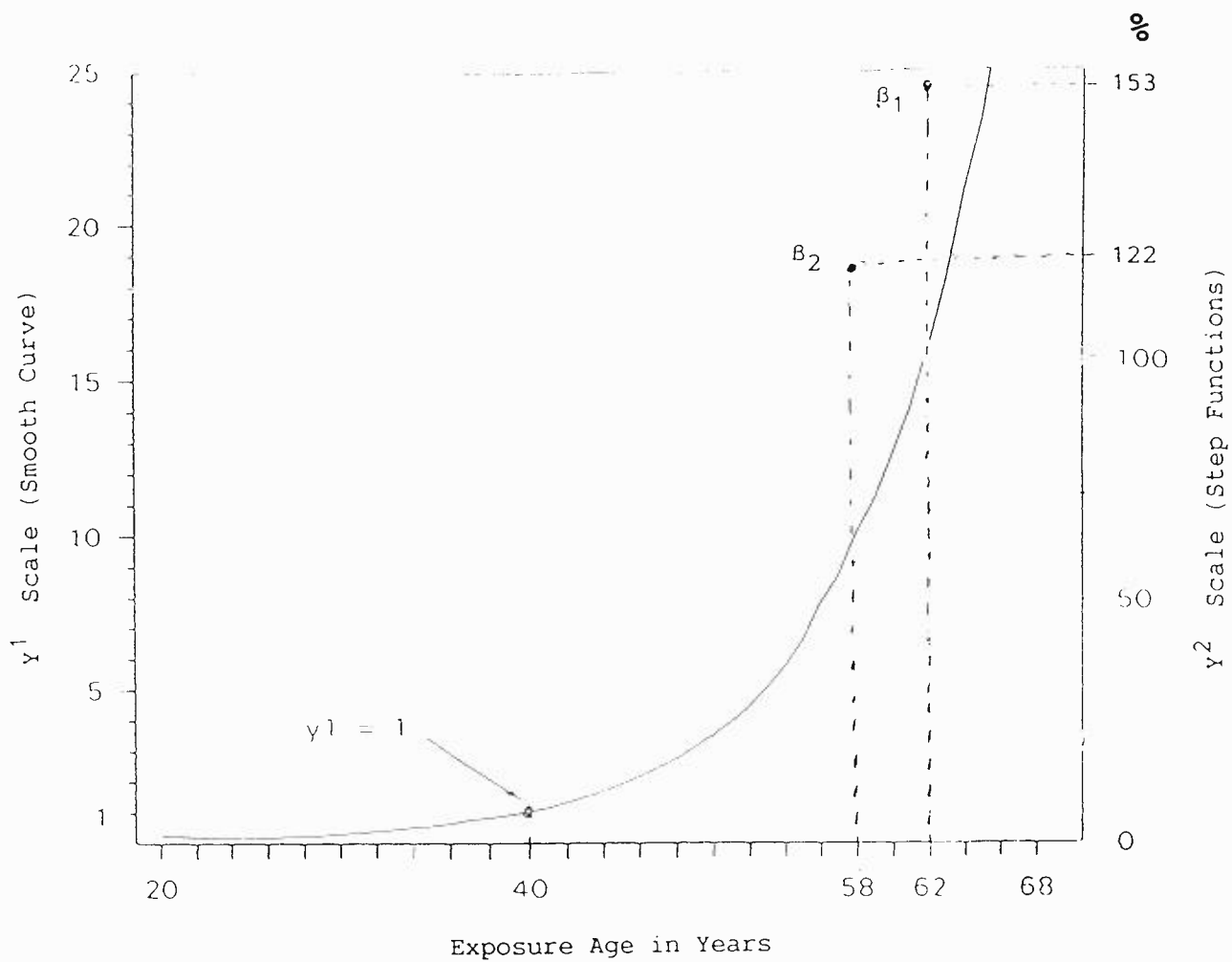
8th Revision	Global only	Global & Local	Local only	Total
140-149	-	2	9	11
150-159*	13	32 (7)	69 (10)	114 (17)
160-163*	25 (2)	38 (8)	63 (16)	126 (26)
170-174*	2	3	3	8
180-189	13	28 (8)	37 (13)	78 (21)
190-199	8 (1)	9 (2)	12 (1)	29 (3)
200-209*	4	13 (4)	14 (2)	31 (6)
Total	65 (3)	125 (29)	207 (42)	397 (74)
*(A Cancers) (1)	No. 44 (2)	86 (19)	149 (28)	279 (49)
	% 67.7	68.8	72.0	70.3

1 = see Kneale *et al*, 1981

Legend to Figures

Figure 1. Age related increase in sensitivity to cancer induction effects of radiation according to several models

Figure 2. Window dose. Model IV estimate for cancer deaths at 75 years



y^1 Scale for exposure age effect (see Kneale *et al*, 1981)

y^2 Scale for percentage increase in risk per 10 mSv (see Tables V and VI)

B_1 Doubling dose for fatal cancers in cohort HXY

B_2 Doubling dose for fatal cancers in cohort H see model IV in Table VI

Fig 1

