Commission of the European Communities

radiation protection

Neutron carcinogenesis

J.J. Broerse (TNO)
G.B. Gerber (CEC)

Directorate-General
Science, Research and Development

1982

EUR 8084 EN

NEUTRON CARCINOGENESIS

= TABLE OF CONTENTS =

ANALYSIS OF TUMOR RATES AND INCIDENCES
- A SURVEY OF CONCEPTS AND METHODS -

A.M.Kellerer and D.Chmelevsky
Institut für Medizinische Strahlenkunde der Universität Würzburg,
Versbacher Str.5, D-8700 Würzburg

1. Introduction

Risk estimates for radiation carcinogenesis are based on observed dose-effect relations. Such relations can be obtained in various ways and, depending on the underlying quantities and the computational procedures, they may differ greatly, even for the same set of experimental data or for the same epidemiological findings.

It is therefore mandatory to achieve coherence in the definition and usage of basic quantities, in the estimates of these quantities, and in the derivation of their standard errors or confidence intervals. It is equally important that suitable tests be utilized for the comparison of incidences or tumor rates in different groups of exposed individuals. Such tests exist for certain types of data but are lacking for others. It is, furthermore, of importance that proper methods be used for the joint analysis of time dependences of tumor rates or tumor prevalences in a number of groups differently exposed. Such methods are always based on implicit assumptions. While the assumptions are unavoidable, they need to be properly stated and to be clearly understood.

In this brief synopsis the definition of basic quantities will be given, the notion of censored data will be explained, estimates and their standard errors will be considered, the existence of non-parametric tests for right-censored data and the lack of non-parametric tests for double-censored data will be discussed. A further point of particular current interest is the joint maximum-likelihood analysis of time and dose dependences of tumor rates or tumor prevalences. For more rigorous and detailed treatments the reader can turn to a number of useful reviews and monographs (1-5).

2. The Simple Case of Uncensored Data

The most simple, although the least common case is that of *uncensored* data. The term uncensored means that the observations are complete, i.e., individuals remain at risk either to the end of the observation period or until the effect of interest occurs. Furthermore it is assumed that one deals with a *manifest disease*, i.e., a disease that is either rapidly lethal or otherwise readily discovered. Examples of manifest diseases are leukemias, osteosarcomas or, among the non-lethal examples, mammary neoplasms in the rat. An example of an *occult disease* - observed in the sacrificed animal, or observed incidentally in animals dead for other reasons - are benign pulmonary neoplasms or, in the case of the rat, also pulmonary malignancies that appear to cause no appreciable life shortening (6). The distinction between manifest and occult tumors is fundamental, because entirely different mathematical procedures are required, that will be considered in sections 3 and 4.

The observation, even of a manifest disease, can be uncensored only in the absence of competing risks, i.e., if no unrelated mortality and no loss from observation for other reason occurs. This is rarely the case. However, the simple example of uncensored data may serve to introduce the basic concepts and quantities.

For the subsequent considerations let $N(t)$ be the number of survivors, i.e. the number of individuals still at risk at time t, and let $n(t)$ be the number of individuals that have incurred the effect up to time t. Let $t=o$ be the time at the start of the experiment and $N=N(o)$ the number of individuals at the time of exposure. Statistical estimates of the quantities will, in the usual way, be designated by a circumflex, and standard errors of the estimates will be included in the formulae.

The most commonly considered quantity is the *cumulative incidence* (or simply *incidence*), $I(t)$, up to time t; it is the probability of an individual to incur the effect up to the specified time. The estimate of $I(t)$ for uncensored data requires no explanation; the standard deviation results from the assumption of statistical independence, i.e., from the assumption that $n(t)$

follows a binomial distribution:

$$\hat{I}(t) = n(t)/N \pm \sqrt{n(t)(1-n(t)/N)} / N \qquad (1)$$

If several tumors can occur, i.e., if animals remain at risk after they have developed neoplasms, one can consider another basic quantity the *mean number of tumors per animal*, R(t). Its estimate for uncensored data is also trivial. It is simply the ratio of the total number, m(t), of tumors up to time t and the number, N, of animals:

$$\hat{R}(t) = m(t)/N \pm \sqrt{m(t)}/N \qquad (2)$$

The standard error in this equation is based on a presumed statistical independence of tumors, i.e., on the Poisson distribution. If the neoplasms are not independent, the standard error in Eq(2) can be too small and one must use the standard error based on the observed numbers of tumors in individual animals. This aspect deserves careful attention, whenever multiple tumors are analysed.

R(t) is also called *cumulative tumor rate*, because it is the integral of the *(differential) tumor rate*, r(t):

$$R(t) = \int_0^t r(\tau)\, d\tau \qquad (3)$$

r(t) is the probability per individual and per unit time (at time t) to develop a tumor. The tumor rate is a fundamental quantity analogous to the mortality rate, and the common term *hazard function* is used in the statistical literature. The age and sex dependent mortality rates and tumor rates in human populations are essential for actuarian epidemiological investigations. However, it is rarely possible in radiation carcinogenesis studies to estimate the time dependence of the differential tumor rates, r(t), as large numbers of observed individuals would be required to achieve reasonable precision. The cumulative rate, R(t), is an integral over the observation period up to time t, and is therefore more readily estimated.

In the subsequent section it will be pointed out that the cumulative tumor rate, R(t), is a meaningful quantity even in those cases where first tumors only are observed.

Cumulative incidence, $I(t)$, and the mean number of tumors per animal, $R(t)$, are related if independence of the individual tumors can be assumed, i.e., if the assumption of a Poisson distribution is valid. $I(t)$ is the probability for at least one tumor up to time t. According to the Poisson statistics $S(t) = \exp(-R(t))$ is the probability for no tumor if the mean number of tumors is $R(t)$. Accordingly one has the following relations between $I(t)$, $R(t)$, and the probability, $S(t)$, for no tumor:

$$I(t) = 1-S(t) = 1-\exp(-R(t)) \quad \text{or} \quad R(t) = -\ln(1-I(t)) = -\ln S(t) \qquad (4)$$

A semi-logarithmic plot of $S(t)$ versus t yields the same curves as a linear plot of $R(t)$. However, it is evident that the numerical estimates from Eqs(1) and (2) need not precisely fulfil relation (4). This is a matter of statistical fluctuations.

An estimation of the cumulative tumor rate, $R(t)$, on the basis of multiple tumors is advantageous because it makes use of the full experimental information. On the other hand, the estimation on the basis of first tumors only avoids the problem of the statistical dependence or independence of tumors.

As with the estimates of the cumulative incidence and the cumulative tumor rate, there are no difficulties with statistical tests for comparison of tumor rates in the simple case of uncensored data. Non-parametric tests are most suitable because they require no assumptions on the time dependence of $I(t)$ or $R(t)$. If only the numbers, $n_1(t)$ and $n_2(t)$, of animals with tumor up to a specified time, t, in the two groups of size N_1 and N_2, are known, the Fisher Exact Probability Test is applicable (see, for example (7)). It is, however, evident that the mere comparison of $n_1(t)$ and $n_2(t)$ at an arbitrary time is not fully satisfactory; even at comparable cumulative incidences the temporal patterns of events in the two groups may differ substantially. It is therefore more efficient to utilize the exact event times and to apply the Mann-Whitney rank-sum test for the comparison of tumor rates in the two groups. If the fraction of individuals affected is small the logrank test or the Breslow test (see section 3.2) can be more suitable.

3. The Common Case of Right-Censored Data

In most experimental and epidemiological investigations one deals with *incomplete observations*; with increasing time, t, the observed sample decreases not merely due to the occurrence of the effect in individuals but also due to other causes such as unrelated mortality, disappearance for other reasons, or incompleteness of observation in some individuals in an ongoing study. For manifest diseases one speaks then of *right-censored* data, and this expresses the fact that one knows for an individual either the actual time, t, of the occurrence of the tumor or one knows that the individual has incurred no tumor up to a time, τ_i, of disappearance due to unrelated reasons.

3.1 *Estimates of R(t) and I(t)*

It is apparent that one can not, for right-censored data, simply apply formulae (1) or (2). If one were to count merely the number of observed tumors or of individuals with tumor and to relate them to the initial number of individuals at risk, one would evidently underestimate the cumulative tumor rate or the cumulative incidence. Surprisingly it is not uncommon that *crude incidences*, are derived from Eq(1) in spite of competing risks that may depend on absorbed dose and on other factors. This may lead to meaningless results, and the use of crude incidences must be strongly discouraged. In too many instances it has invalidated the results of otherwise valuable studies.

Proper competing risk corrected estimates are comparatively simple and have been in use (as documented by Groer (8)) for centuries. No involved statistical theory is needed to understand the proper estimate of the cumulative tumor rate, R(t), for right-censored data. One merely needs to sum, up to the specified time, t, all quotients of newly arising tumors and individuals, N(t), still at risk. In the simple case where all tumors are observed individually at their occurrence times, t_i, the quotients are $1/N(t_i)$ and the estimate is therefore:

$$\hat{R}(t) = \sum \frac{1}{N(t_i)} \pm \sqrt{\sum \frac{1}{N(t_i)^2}} \qquad (5)$$

The standard deviation results from the properties of the Poisson distribution (see (9)). The summations extend over all event times, t_i, up to time t.

If the observation is not continuous but at distinct times, and if k_i events are found at time t_i, the term $1/N(t_i)$ is replaced by k_i/N_i, and $1/N(t_i)^2$ is replaced by k_i/N_i^2, where N_i is a mean number of animals at risk during the intervening interval of no observation. Analogous modifications apply to Eq(7).

Eq(5) can be applied to first tumors, but it can equally be applied if successive tumors can occur in one animal (9). This requires no different formula because animals with a previous tumor remain then at risk and continue to be included in N(t), until they are removed due to competing risks. It is important to appreciate this point: inclusion of multiple tumors increases the number of terms in the summation, but it decreases the magnitude of the individual terms due to the larger values $N(t_i)$. The estimates obtained in the two ways are equivalent, if the tumors are statistically independent. If they are not, i.e., if occurrence of a tumor increases the probability for subsequent tumors, or if animals have inherently different tumor rates, $\hat{R}(t)$ computed from all tumors will exceed $\hat{R}(t)$ computed from first tumors only. It must also be noted that the standard error in Eq(5) is then too small. The term (*actuarian*) *mean number of tumors* is synonymous with cumulative tumor rate if R(t) is based not only on first tumors.

If Poisson statistics can be assumed, or if first tumors only are utilized, the estimate from Eq(5) can with Eq(4) also provide the value of the cumulative incidence:

$$\hat{I}(t) = 1 - \exp(-\hat{R}(t)) \qquad (6)$$

This estimate has certain advantages. However, another, largely equivalent estimate has been far more common. Although it has a much earlier history, it is usually called the Kaplan-Meier *product-limit estimate*. In the same way as the *sum-limit estimate* of Eq(5) it is here given in its simple form for individually resolved event times:

$$\hat{I}(t) = \Pi\left(1 - \frac{1}{N(t_i)}\right) \pm \left(1 - \hat{I}(t)\right) \cdot \sqrt{\sum \frac{1}{N(t_i)^2}} \qquad (7)$$

As in Eq(5) the summations extend over all times, t_i, up to time t.
The expression for the standard error is called the Greenwood formula. Both the estimate and the standard error lead numerically to nearly the same values

as Eq(6) with Eq(5), provided the number of individuals at risk is not too small. It is therefore arbitrary whether one chooses the sum-limit estimate or the product-limit estimate to obtain the cumulative incidence (or *actuarian incidence*).

A merely technical point deserves consideration. The estimates of Eq(5) to (7) provide step functions. The discrete steps have evidently no biological meaning and they can make it difficult to read graphs with a set of intersecting curves. It is therefore advisable to draw the resulting curves as polygons connecting the midpoints on the vertical steps.

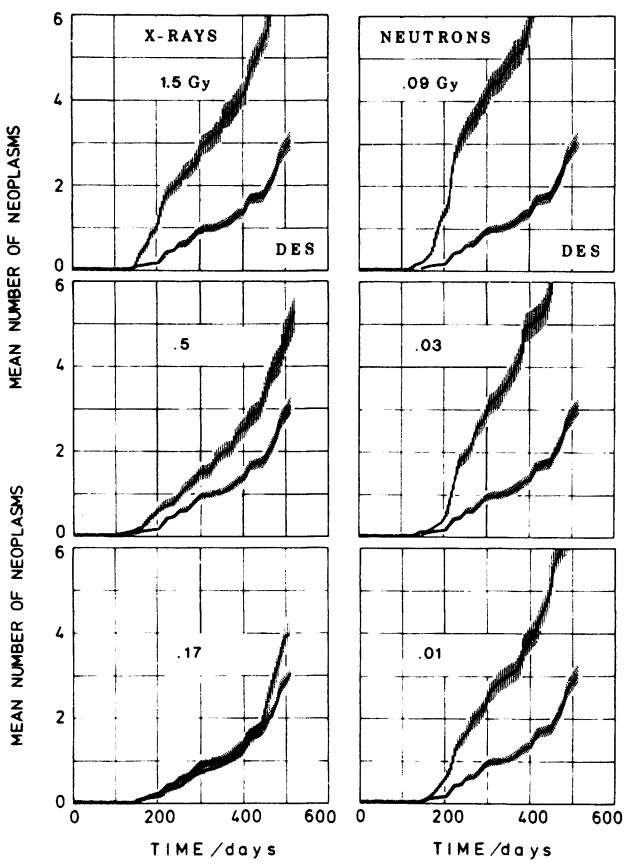

Figure 1.

Cumulative tumor rates, R(t), i.e. actuarian mean numbers of adenocarcinomas, in DES treated ACI rats exposed to x-rays and to 430 keV neutrons in a recent experiment by Shellabarger et al. (10).

The shaded areas are standard deviations according to Eq.(5). The lower curves represent the control group.

Fig.1 gives results from a recent experiment of Shellabarger et al.(10). For female ACI rats treated with DES the actuarian mean number, R(t), of mammary adenocarcinomas is given as a function of the dose of 430 keV neutrons and of x-rays.

3.2 *Tests for Right-Censored Data*

Neither the Exact Probability Test nor the Mann-Whitney rank-sum test are applicable to right-censored data. However, the more recently introduced *Mantel-Haenzel* (or *log-rank*) *test* or certain closely related non-parametric tests such as the *Breslow test* (for detailed explanations see (1,2)) can be utilized.

The log-rank test can be understood in the same way as Eq(5). Event rates in two groups are compared. Whenever an event occurs one considers the probability, under the null-hypothesis, that it be in group 1. This probability is equal to the relative proportion of individuals in group 1 at the time t_i when the event occurs $p_i = N_1(t_i)/(N_1(t_i)+N_2(t_i))$. If $t_i (i = 1,2 \ldots I)$ are the event times in the two groups, the expected number of events, $\bar{n}_1$, in group 1 and the corresponding variance, σ_1^2, are:

$$\bar{n}_1 = \sum_{i=1}^{I} p_i \qquad \text{and} \qquad \sigma_1^2 = \sum_{i=1}^{I} p_i\,(1-p_i) \tag{8}$$

The variance results from the binominal distribution and the presumed independence of results in successive observations. The difference between the calculated expectation, $\bar{n}_1$, and the actually observed number, n_1, of events in group 1 can then serve as test statistic to assess the acceptability of the null-hypothesis of equality of the tumor rates in the two groups. Accordingly one uses the test statistic:

$$z = (|n_1 - \bar{n}_1| - 0.5) \,/\, \sigma_1 \tag{9}$$

the value of z is tested against the standard normal distribution, i.e., if z exceeds 1.96 equality of the tumor rates in the two groups is rejected with two-sided error probability 0.05.

The term -0.5, the so-called continuity correction, in Eq(9) is essential for small samples; failure to include the continuity correction can lead to actual error levels that are substantially in excess of the assumed error levels (11). An explicit form of the test (12), not based on the standard normal distribution, has been applied to the study of neutron induced mammary neoplasms in the rat (9).

Peto and Peto (13) have proposed generalizations of the log-rank test that result when different weight factors, w_i, are introduced in Eqs(8). Of particular importance is the possibility to assign larger weights to the earlier observations; for example one can use descending ranks with successive events (12). The important Breslow test results when the numbers of individuals still at risk in both groups are chosen as weight factors. The Breslow test has more power than the simple log-rank test, if tumor rates in the two groups differ only in a initial phase of the observation period. If tumor rates differ by the same factor throughout the observation period (see section 3.3 for the proportional hazards model) the log-rank test has the highest power (13).

3.3 *Joint Estimates from Several Observed Groups*

Radiation-carcinogenesis studies involve frequently a set of groups exposed differently. As pointed out, there are then methods to estimate the time dependences and standard errors of the cumulative incidence or the cumulative tumor rate in the individual groups. In a large experiment where the time dependences in the individual groups are well defined, the results may be sufficient to achieve a coherent overall picture of the time and dose response. An example is Shallabarger's earlier study (9) on mammary neoplasms in neutron exposed Sprague-Dawley rats. However, in the more common case of small or moderately sized groups the statistical fluctuations may be substantial and it may be difficult to recognize a coherent picture of the time and dose dependence. More complex methods are then required to arrive at consistent and comparable results. Such methods must always be based on certain implicit assumptions, and this introduces an element of arbitrariness. While this is unavoidable, the assumptions need to be understood and need to be pointed out clearly.

The *proportional hazards model* and its solution by a method of Cox is virtually the only approach for simultaneous fits utilized in radiation-carcinogenesis studies. It will therefore be briefly considered. However, other possibilities to obtain a joint maximum likelihood fit of R(t) or I(t) will also be indicated. All treatments are based on the assumption that the differential or integral tumor rate has a certain dependence on time after exposure, and that this dependence changes in a relatively simple way with increasing absorbed

dose or with the variation of other factors. The proportional hazards model postulates that the tumor rates are increased by a dose-dependent but time independent factor:

$$R(t,D) = \lambda(D) \cdot R(t) \qquad (10)$$

where R(t) is a common *base-line function*. It is the objective of the computational procedure to obtain best estimates of the base-line function and of the proportional hazard factors, $\lambda(D_i)$, for the individual groups exposed to doses D_i. The problem is non-parametric in the sense that one utilizes no analytical expression for the base line function R(t); instead one searches generally for the best solution that need merely be monotonic in t.

Kalbfleisch and Prentice (5) give excellent accounts of the Cox model and of related topics; in the present context it is therefore sufficient to appreciate the essential problem. This is to find that base-line function, R(t), and those parameters, λ_i, for the individual groups that maximize the *likelihood* for the observed result. The likelihood is computed under acceptance of the observed event times, t_i (i = 1,2 .. I), and censoring times, τ_k (k = 1,2 .. K). For simplicity it will again be assumed that each event time correspond to one new tumor and each censoring time corresponds to the loss of one individual for unrelated reasons. One can then readily show that the likelihood is equal to:

$$L = \prod_l i(t_l) \quad \prod_k (1-I(\tau_k)) \qquad (11)$$

where I(t) is the cumulative incidence, and i(t) is the derivative of I(t).[+)] The products extend over all times, t_l, of the occurrence of a tumor and all times, τ_k, of censoring (end of observation of an individual without tumor).

The solution of the proportional hazards model for right-censored data requires only a relatively simple iterative optimization algorithm. This appears to be the reason that the few derivations of joint maximum likelihood solutions in radiation-carcinogenesis studies have exclusively used the proportional hazards model (see for example (14)(15)). One must, however, realize that there are

---

[+)] In certain numerical computations, for example the Cox solution, I(t) is represented by a function with discrete steps, and $i(t_l)$ is then not the derivative of $I(t_l)$, but is equal to its change at t_l.

alternatives. With the further sophistication and extension of experimental investigations or epidemiological studies it will be essential to explore and compare different treatments and to assess their possible bias. Two main alternatives to the proportional hazards model will therefore be mentioned. It is evident, that there are others.

One alternative to the proportional hazards model is the *time-shift model*. It postulates that the tumor rates remain unchanged, but are shifted forward in time with increasing dose:

$$R(t,D) = R(\, t+s(D)\,) \;, \quad \text{i.e.} \quad I(t,D) = I(\, t+s(D)\,) \tag{12}$$

Shellabarger's extensive investigation of mammary neoplasms in the Sprague-Dawley rat (9) has led to results that are consistent with this relation. However, no joint maximum likelihood solution has yet been derived for these results or similar sets of data.

A second alternative is the *accelerated time model*. This model has found applications in the field of industrial testing procedures; it is there known as the *accelerated failure-time model* (16). It invokes an acceleration factor,not a shift in time:

$$R(t,D) = R(\, \alpha(D)\cdot t\,) \;, \quad \text{i.e.} \quad I(t,D) = I(\, \alpha(D)\cdot t\,) \tag{13}$$

The accelerated time model has not been utilized in radiation-carcinogenesis studies with right-censored data. However, it may deserve as much attention as the proportional hazards model.

A judgement of the comparative applicability of different models may not be possible, except in large scale investigations. Furthermore the proportional hazards model and the accelerated time model are equivalent if the baseline function is a simple power of time:

$$R(t,D) = \lambda(D)\; t^{p} = (\alpha(D)\cdot t)^{p} \;, \quad \text{with} \quad \alpha(D) = \lambda(D)^{1/p} \tag{14}$$

For this *Weibull model* the question is meaningless whether irradiation leads to *more* tumors or to the same number of tumors *at earlier times*.

The comparatively simple Eq(14) is an example of a parametric model. It is evident that there are numerous other possibilities. The more general Weibull model:

$$R(t,D) = \lambda\ (t+s)^p \qquad (15)$$

can be used in a variety of ways depending on the assumed dose dependence of one or of more of its three parameters λ,s, and p. A different parametric model, not infrequently invoked (see for example (18,19)) is the assumption that I(t,D) follows a cumulative log-normal distribution.

Parametric models may be useful for the analysis of the time and dose dependence of tumor rates. In general, however, it will be better to avoid assumptions that may be too narrow, and to perform the analysis with non-parametric base-line functions. Regardless of the choice of models, it is essential to obtain the simultaneous maximum likelihood solutions by computer. With the non-parametric approach the computer solution is unavoidable, although it is somewhat complex. With the parametric models the computer solutions happen to be fairly simple (see (3)). Visual fits are uncertain and easily misleading because one tends to estimate the base-line function from the most informative group in the experiment with mere subsequent adjustment of one or two parameters to the remaining groups. In the rigorous solution the shape of the base-line function is jointly determined by the entirety of data.

Lafuma and colleagues investigate in their current study (20,21) the full spectrum of neoplasms in Sprague-Dawley rats exposed to fission neutrons. In an initial step of the analysis, with some low dose groups and the controls still unfinished, all lethal carcinomas and sarcomas have been pooled. Fig.2 gives, as an example, the sum-limit estimates (see Eq(5)). To keep this common graph readable the standard errors are omitted. Due to the moderate size of the groups exposed to different doses there are fluctuations that make it difficult to judge the precise trend of the time and dose dependence. The Cox proportional hazards model has, therefore, been applied to obtain a more coherent picture, and the results are given in Fig.3. It is evident that this type of analysis facilitates the derivation of dose-effect relations or RBE-dose relations. Only preliminary results are available from this current investigation. However, it is evident that a comparative study of the different models will be required.

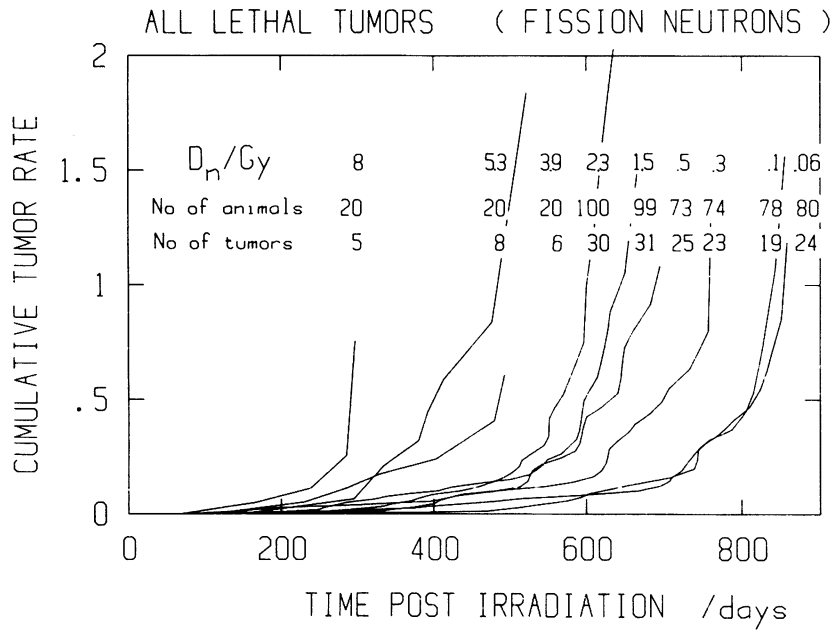

Figure 2. Cumulative tumor rates, R(t), for all lethal carcinomas and sarcomas in Sprague-Dawley rats exposed to fission neutrons in a current experiment by Lafuma et al.(20,21). The curves are individual estimates according to Eq.(5) for the dose groups with completed observation; the standard deviations are omitted to keep the graph readable.

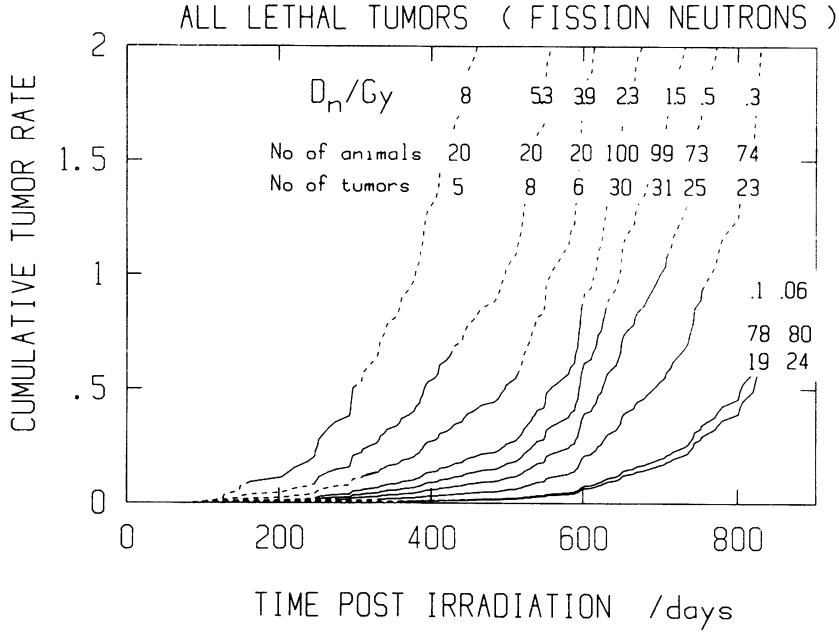

Figure 3. Analogous graph to Fig.2, with results based on the proportional hazards model. The curves are given as broken lines where there are less than 4 animals dead or less than 4 animals alive in the group.

Fig.4 exemplifies a slightly different and not uncommon representation that is essentially equivalent to Fig.3. This graph gives the probability, $S(t)=\exp(-R(t))$, for no tumor at time t.

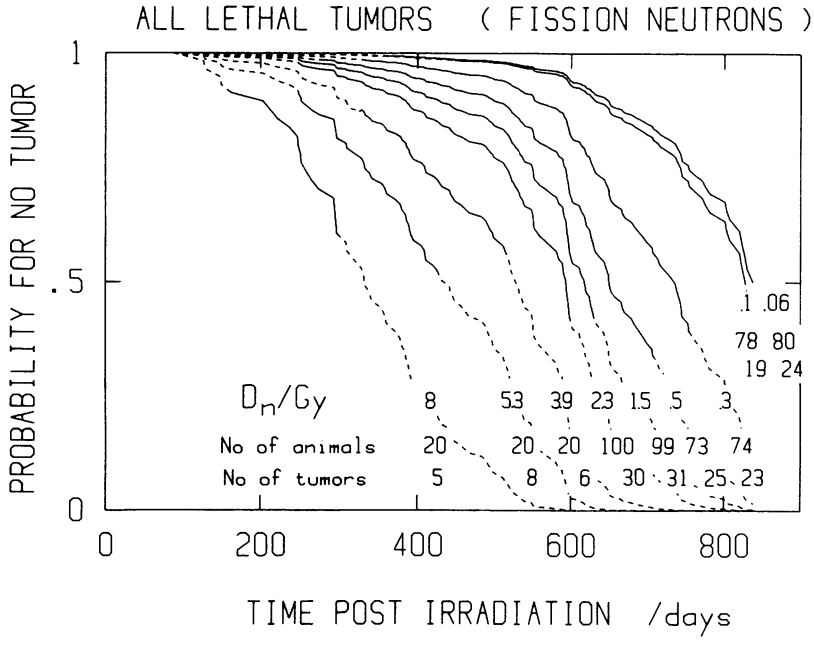

Figure 4.

The same results as in Fig.3, plotted as probability for no tumor, $S(t) = \exp(-R(t))$.

4. The Case of Double-Censored Data

The term *double-censored data* refers to the case of *occult* diseases that are observed only incidentally in sacrificed animals or in individuals dead for other reasons. In the present context it is assumed that the disease is entirely non-fatal. Even with this assumption the statistical problems in the analysis are considerable. One must, furthermore, note that there is a large gray area of diseases that are neither fatal nor entirely non-fatal; additional complications will then arise that will not be considered.

With double censored data one never knows actual times of occurrence of a tumor. When a dead animal bears the tumor, the event time, t_i, is *smaller* than the time, τ_i, of death. When the dead animal bears no tumor the hypothetical event time, t_i, is *larger* than τ_i.

4.1 *Estimate of I(t)*

The basic quantity to be investigated is the probability of an animal, at time t, to bear the tumor. This corresponds to the cumulative incidence, I(t),

discussed earlier, but the term *prevalence* is more commonly used.[+)] Estimation of I(t) and its standard error is simple in an experiment with serial sacrifices; Eqs(1) and (2) are then applicable. However, experiments with serial sacrifices require large numbers of animals. If a multiplicity of absorbed doses or other factors are investigated the approach may become quickly prohibitive. Survival experiments, or survival experiments combined with sacrifices, can then be more economical. However, comparatively little use is made of this possibility, this being apparently due to the statistical problems arising in the analysis. On the other hand, Hoel and Walburg (22) have, a decade ago, exemplified and recommended the method that is presently known as *isotonic regression* and that can for double-censored data be considered the analogon of the product-limit estimate for right-censored data. Isotonic regression is a relatively simple algorithm that provides, with the constraint of monotonicity, a maximum likelihood estimate of the prevalence I(t) of an occult disease (23).

Apart from Hoel and Walburg's work and a recent study (6), it is difficult to find examples where isotonic estimates have been used for the analysis of radiation carcinogenesis. In view of its unfamiliarity the algorithm is, therefore, explained in the appendix. The somewhat difficult problem of the derivation of standard errors for isotonic regression will not be considered, as there appears to be no valid theoretical treatment of the topic.

Fig.5 gives, as an example for isotonic regression, the estimates of lung cancer prevalence in Sprague-Dawley rats exposed to fission neutrons. As in Figs.2 and 3 the data are from the on-going experiment by Lafuma et al.(20,21). The isotonic regression curves are, in actuality, step functions. However, as with the estimates for right-censored data, it is convenient to connect the midpoints of the vertical steps to obtain better readability of the graph. The initial part of the first step is retained to indicate more clearly the first death of an animal with tumor.

[+)] For partially lethal tumors the concept of prevalence differs from that of cumulative incidence, as it depends on the mean survival time.

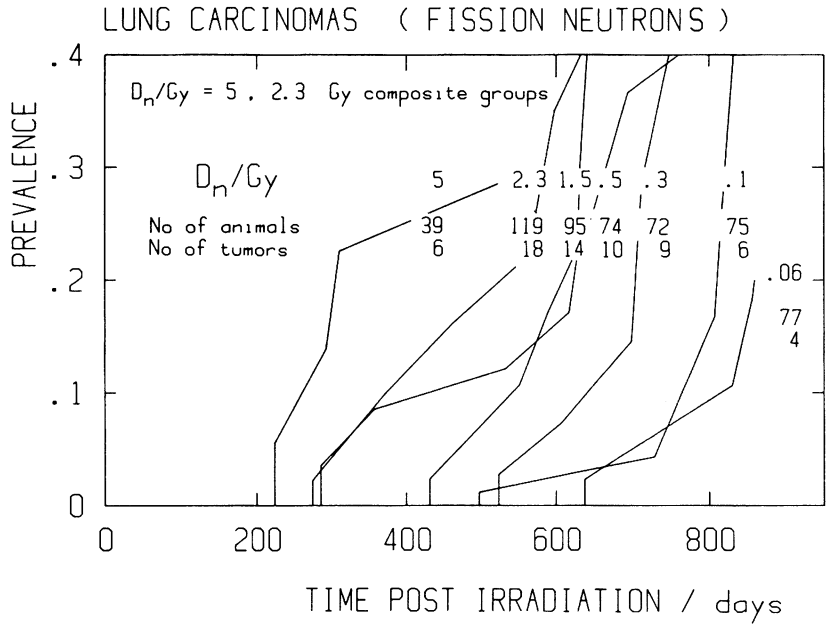

Figure 5. Prevalence of lung carcinomas in fission neutron irradiated Sprague-Dawley rats from a current experiment by Lafuma et al.(20,21). The curves are separate isotonic regression estimates for the dose groups with completed observation. The control incidence of lung cancer is small (≃.003 life-time incidence).

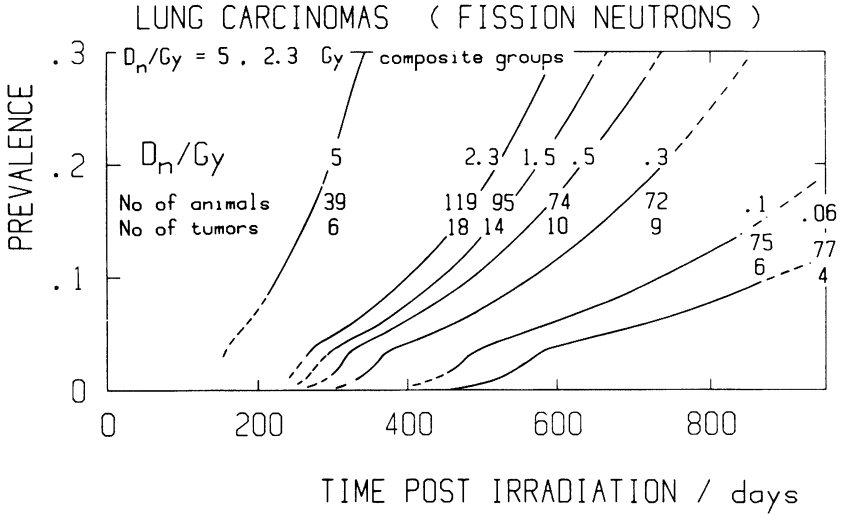

Figure 6. Analogous graph to Fig.5 with a joint maximum likelihood solution based on the accelerated time model.

4.2 *Tests for Double-Censored Data*

It would be very desirable to obtain a non-parametric test for the comparison of prevalences in two groups with double-censored data, particularly if such a test were valid for different degrees of censoring in the two groups. No analogon to the log-rank or to the Breslow test for double-censored data appears to exist, at present. For sufficiently large samples suitable test procedures can, of course, be established; but the development of a generally applicable test remains a challenge of particular interest.

4.3 *Joint Maximum Likelihood Estimates of the Prevalence*

As with right-censored data, one will frequently encounter a situation where the time dependence of the prevalence is to be estimated in a number of groups differently exposed. Isotonic regression, applied separately to the groups, will then be helpful but need not always provide a sufficiently consistent picture of the time and dose dependence. The data in Fig.5 illustrate the point. The arguments in section 3.3 apply equally to double-censored data. One deals with the same problem of estimating a common base-line function for the prevalence, $I(t)$, and individual parameters, according to an assumed model, for the individual groups. The likelihood has the form:

$$L = \prod_l I(t_l) \cdot \prod_k (1-I(\tau_k)) \qquad (16)$$

The products extend over all death times, t_l, of animals with the tumor and all death times, τ_k, of animals without the tumor. The computational details are somewhat different, but the same essential models, expressed by Eqs(10, 12,13) can be utilized, i.e., possible approaches are again the proportional hazards model, the time-shift model, and the accelerated time model.

The determination of the maximum likelihood solutions, i.e. of the best base-line function, $I(t)$, and the best set of parameters, $\lambda(D_i)$, $s(D_i)$, or $a(D_i)$ (see Eqs(10,12,13)), requires relatively powerful non-linear optimization procedures (see for example (24)), and convergence of the solution is not

always without problem. Such complications are, however, in no way commensurable with the over-all efforts invested in radiation-carcinogenesis experiments. There is, therefore, little justification for making inadequate use of available numerical methods in the analysis of the experiments. Details of the computational procedures are described in a recent analysis of radon inhalation studies with Sprague-Dawley rats by Lafuma et al.(6).

The same results from the current fission-neutron experiments that have been utilized for Fig.5, can serve to exemplify the joint maximum likelihood fit. The solutions in Fig.6 are obtained for the accelerated time model. It is apparent that these results provide a more consistent picture of the time and dose dependence of the prevalence of pulmonary carcinomas in the fission neutron exposed animals. These preliminary and incomplete results from an on-going study are given for illustration. It is evident that evaluation of the ultimate results in terms of all three basic models will be required.

Parametric maximum likelihood fits will not be considered because there is rarely enough readily recognizable evidence in double-censored observations to justify the *a priori* choice of a particular analytical expression for the prevalence. A more reliable approach is, therefore, the non-parametric analysis that identifies the best solution among a much broader spectrum of possible dependences of prevalence on time. The adoption of a parametric model may, however, be an eventual second step after an exploratory investigation by non-parametric methods.

5. Concluding Remarks

The subsequent table indicates some of the statistical concepts and tools that have been considered. It is evident that there are various complexities and problems that have not been treated.

The emphasis has been on the need to employ available estimates, tests, and maximum likelihood algorithms. While competing risk corrected methods are essential, it must be kept in mind, that they have limitations. The most serious limitation can arise from a lack of statistical independence between the observed effect and the competing risks. When interdependence exists, difficult statistical problems result; they are encountered generally in the case of

partially lethal diseases. There are no fully adequate solutions, but the existing reviews (1-5) will be helpful even in these cases.

TYPE OF DATA	NATURE OF OBSERVATION	COMPETING RISKS	ESTIMATES	TESTS	JOINT MAXIMUM LIKELIHOOD FIT	
UNCENSORED	MANIFEST	NO	TRIVIAL	U-TEST	PROPORTIONAL HAZARDS MODEL	PROPORTIONAL HAZARDS MODEL TIME-SHIFT MODEL ACCELERATED TIME MODEL
RIGHT-CENSORED	MANIFEST	YES	I(t): PRODUCT-LIMIT ESTIMATE R(t): SUM-LIMIT ESTIMATE	LOG-RANK TEST BRESLOW TEST	(SIMPLE OPTIMIZATION ROUTINE FOR COX SOLUTION)	
DOUBLE-CENSORED	OCCULT	YES	I(t): ISOTONIC REGRESSION			(COMPLEX OPTIMIZATION ROUTINES)

Table 1 Synopsis of Concepts and Methods

APPENDIX

The Algorithm for Isotonic Regression:

The subsequent explanation follows closely the more detailed treatment of Barlow et al.(23).

Let τ_i be the times where deaths occur, that may or may not be sacrifices. The number of animals that die at τ_i is n_i, the number with tumor m_i. Furthermore let N_i be the total number of deaths, and M_i be the total number of animals dead with tumor up to and including time τ_i. All n_i are equal to 1, if the times of deaths are individually resolved.

The following table gives a simple fictitious data set with 13 deaths. As with the sum-limit or product-limit estimates the numerical values of the time do not enter the actual calculations. In Fig.A1 the total number, M_i, of animals dead with tumor is plotted versus the total number, N_i, of animals dead.

time (days)	No of animals dying with tumor	No of animals dying	crude estimate of prevalence	isotonic estimate
230	0	1	0	0
260	1	1	1	.25
290	0	1	0	.25
310	0	1	0	.25
350	0	1	0	.25
380	2	3	.66	.57
560	2	2	1	.57
610	0	2	0	.57
650	1	1	1	1

Table 2 Data set

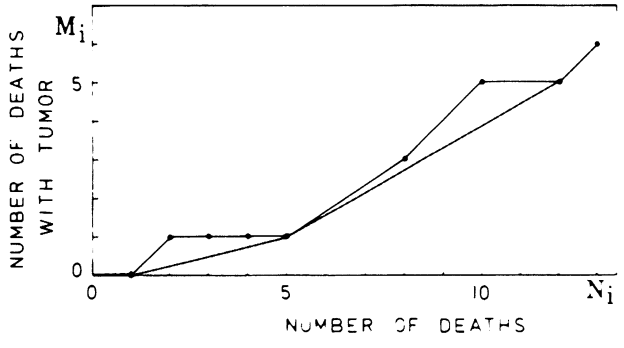

Figure A1.

Illustration of the algorithm for isotonic regression

Each dot corresponds to a time, τ_i, and the slope of the line segment to the preceding dot is equal to the naive estimate, m_i/n_i, of the prevalence at the time τ_i. The naive estimates do not fulfill the required condition that the prevalence be an increasing function of time.

To obtain the isotonic regression, an irregular pair of adjacent segments, i.e. one with decreasing slopes, is replaced by a single segment joining the end points. If further irregular pairs are left, the procedure is repeated. In a computer routine one may go from left to right, and repeat the procedure until full regularity is achieved. It is evident from Fig.A1, that the final result is the shortest convex curve (taut string) below the dots.

The slopes of the resulting curve (broken line in Fig.A1) are the isotonic estimates of the prevalence at the corresponding times. The inflections are still assigned the preceding slope.

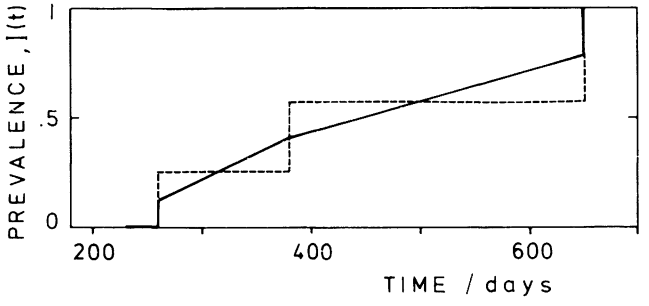

Figure A2.

Isotonic regression for the data from Table 2.

The estimated prevalences are given in Fig.A2 as a step function (broken line). For easier readability of a graph with several curves the step function may be replaced by a polygon (solid line) interpolating the steps. In Fig.A2, as in the earlier Fig.5, the midpoints of the vertical steps have been connected. This is an *ad-hoc* convention that need not be the optimal procedure for isotonic estimates.

References

1. Peto, R., Pike, M.C., Armitage, P., Breslow, N.E., Cox, D.R., Howard, S.V., Mantel, N., McPherson, K., Peto, J., Smith, P.G.: Design and Analysis of Randomized Clinical Trials Requiring Prolonged Observation of Each Patient. I.Introduction and design. Br.J.Cancer 34, 585-612, 1976 II.Analysis and examples. Br.J.Cancer 35, 1-39, 1977.

2. Breslow, N.: Analysis of Censored Survival Data. Proc.23rd Conference on the Design of Experiments in Army Research, Development and Testing at the Naval Postgraduate School in Monterrey, October, 1977.

3. Kalbfleisch, J.D., Prentice, R.L.: The Statistical Analysis of Failure Time Data. Wiley, New-York, 1980.

4. Peto, R., Pike, M.C., Day, N.E., Gray, R.G., Lee, P.N., Parish, S., Peot, J., Richards, S., Wahrendorf, J.: Guidelines for Simple, Sensitive Significance Tests for Carcinogenic Effects in Long-Term Animal Experiments. p.311-423 Annex to IARC (1980) IARC Monographs on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Suppl.2, Long-term and short-term screening assays for carcinogens: a critical appraisal, Lyon, IARC.

5. Peto, R.: Guidelines on the Analysis of Tumour Rates and Death Rates in Experimental Animals. Br.J.Cancer 29, 101-105, 1974

6. Chmelevsky, D., Kellerer, A.M., Lafuma, J., Chameaud, J.: Maximum Likelihood Estimation of the Prevalence of Non-Lethal Neoplasms - An Application to Radon Daughter Inhalation Studies- Radiat.Res., to appear June 1982.

7. Siegel, S.: Nonparametric Statistics. McGraw-Hill Book Company, New York-Toronto-London, 1956.

8. Groer, P.G.: Competing Risk Theory and Radiation Risk Assessment. In: Proc.5th Int.Congress of IRPA, Vol.I, p.231-234, 1980.

9. Shellabarger, C.J., Chmelevsky, D., Kellerer, A.M.: Induction of Mammary Neoplasms in the Sprague-Dawley Rat by 430-keV Neutrons and X-Rays. J.Natl.Cancer Inst. 64, 821-833, 1980.

10. Shellabarger, C.J., Chmelevsky, D., Kellerer, A.M., Stone, J.P., Holtzman, S.: Induction of Mammary Neoplasms in the ACI Rat by 430 keV Neutrons, X-Rays, and DES. In preparation.

11. Kellerer, A.M., Chmelevsky, D.: Small Sample Properties of Censored Data Rank Tests. Biometrics, in press

12. Kellerer, A.M.: Comparison of Effect Rates in Samples of Time Decreasing Size. In: Annual Report on Research Project, Rad.Res. Lab.Columbia University, p.214-222, USAEC -COO-3243-2, Springfield, 1973.

13. Peto, R. Peto, J.: Asymptotically Efficient Rank Invariant Test Procedures. J.R.Statist.Soc.A, 135, Part 2, 185-206, 1972.

14. Prentice, R.L. Peterson, A.V., Marek, P.: Dose-Mortality Relationships in RFM Mice following ^{137}Cs γ-Ray Irradiation. Radiat.Res.90, 57-76 (1982).

15. Peterson, A.V., Prentice, R.L., Marek, P.: Relationship between Dose of Injected ^{239}Pu and Bone Sarcoma Mortality in Young Adult Beagles. Radiat.Res.90, 77-89 (1982),

(see also: Peterson Jr., A.V., Prentice, R.L., Ishimaru, T., Kato, H., Mason, M.: Investigation of Circular Asymmetry of Cancer Mortality of Hiroshima and Nagasaki A-Bomb-Survivors, to appear in Radiat.Res.).

16. Cox, D.R.: Regression Models and Life Tables (with discussion). J.R.Stat.Soc.B,34, 187-220, 1972.

17. Peto, R., Lee, P.: Weibull Distributions for Continuous-Carcinogenesis Experiments. Biometrics 29, 457-470, 1973.

18. Gössner, W., Luz, A.: Biological Factors as Illustrated in Work with Short-Lived α-Emitters. In: Bone and Bone Seeking Radionuclides - Physiology, Dosimetry, and Effect, p.137-153 Harwood Academic Publishers, 1981.

19. Johnson, N.L., Kotz, S.: Distributions in Statistics. Continuous Univariate Distributions, Vol.1,2, Boston: Houghton Mifflin, 1970.

20. Lafuma, J., Morin, M., Masse, R.: Cancer Induction in Rats After Fission Neutron Irradiation with Special Emphasis on Lung Cancers. Proc.'European Seminar on Neutron Carcinogenesis', Rijswijk, 1982.

21. Lafuma, J., personal communication.

22. Hoel, D.G., Walburg Jr, H.E.: Statistical Analysis of Survival Experiments. J.Natl.Cancer Inst.49, 361-372, 1972.

23. Barlow, R.E., Bartholomew, D.J., Bremner, J.M., Brunk, H.D.: Statistical Inference under Order Restrictions; The Theory and Application of Isotonic Regression, New-York: John Wiley & Sons, 1972.

24. Himmelblau, D.M.: Applied Nonlinear Programming. McGraw-Hill, New York, 1972.