[CANCER RESEARCH 46, 4372–4378, September 1986]

Age-specific Models of Mortality and Tumor Onset for Historical Control Animals in the National Toxicology Program's Carcinogenicity Experiments

Christopher J. Portier, John C. Hedges, and David G. Hoel

National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709 [C. J. P., D. G. H.], and 3398 Columbia Drive, Pittsburgh, Pennsylvania 15234 [J. C. H.]

ABSTRACT

This paper models general survival and the distribution of tumor onset times for various tumors in the data base of control animals developed by the National Toxicology Program. For general survival, a modified Weibull model is shown to give an adequate fit for both Fischer 344 rats and C57BL/6 × C3H F_1 mice. In addition, data from control animals in a lifetime study of asbestos are used to support the extension of these survival curves beyond 2 years in Fischer rats. The distributions of tumor onset times are modeled using a two-parameter Weibull model. For many common tumor types, this model yielded a very good fit to the data. Finally, a summary measure of the contribution of a tumor to mortality is given.

INTRODUCTION

Lifetime rodent tumorigenicity experiments are utilized to assess the qualitative and quantitative carcinogenic response to chemicals. In order to estimate the risk of cancer associated with a certain level of exposure, the relationship between dose and response must be estimated (1). This risk can depend on the temporal aspects of exposure as well as the degree of exposure. Thus, for some tumors the models used to estimate the cancer risk must provide for both of these effects (2, 3).

This report presents statistical models which reasonably predict general survival and the distribution of tumor onset times in untreated animals. It is realized that in virtually every study the concurrent control group is the most appropriate control group to use in determining cancer risk. However, the models to be presented in this paper can be very useful in the design and analysis of tumorigenicity data. Other authors (4) have given several uses for the historical control data which include testing for increased risk in the case of rare tumors and quality control. The model estimates presented in this paper can be used in similar ways. Also, this information can be useful in designing the experiment by giving estimates of the proportion of animals with a tumor as a function of time. Design considerations related to these estimates would include study duration and interim sacrifice times. Finally, simulation procedures used to understand the methods of risk assessment (*e.g.*, Refs. 5 and 6) can be improved by basing their assumptions upon the historical information provided by this research.

MATERIALS AND METHODS

Data

General Survival. In recent years, the NTP[1] has conducted a large number of carcinogenesis experiments, commonly using male and female Fischer 344 rats and male and female C57BL/6 × C3H F_1 (hereafter called B6C3F_1) mice. After considering many factors, the NTP established a historical control data base. This data base consists of all chemicals for which Technical Reports are greater than 193 and for which the laboratory's pathology diagnoses were approved by March 1983. For male and female Fischer rats, this data base consists of 2320 and 2370 animals, respectively, from the untreated control groups of 47 studies. For B6C3F_1 mice, there are 2343 males and 2486 females from 51 studies. Most control groups consisted of 50 animals and all are from 2-year studies. The magnitude of this data base prevents the inclusion of individual animal records in this paper. However, published summaries of tumor counts, number surviving to sacrifice, etc., are available (see, *e.g.*, Ref. 4). In addition, interested readers should be able to obtain more detailed information regarding the data base from the National Toxicology Program. We will refer to this data base as the NTP historical control data base. Using this data base, we will discuss models which reasonably predict the general survival and the distribution of tumor onset times in these control animals.

Tumor Onset. There are 10 sacrifice times in the NTP historical control data base for Fischer rats. The first sacrifices occur at 109 weeks of age and continue weekly up through age 118. For mice, there are only 8 sacrifice times starting at week 109 and continuing to week 116. This is sufficient information to apply the methods being used here; however, because all of the sacrifice information is at the end of the observation period, the model estimates will be extremely variable for early times. To improve the precision of these model estimates, we have added the untreated control groups from 7 other National Toxicology Program studies to the historical control data base. These data are summarized in Table 1. Except for butyl benzyl phthalate, the studies to which these groups belong were gavage experiments which used a vehicle control and an untreated control. The untreated control groups were found to be unnecessary after the study was begun, so the animals were sacrificed. For butyl benzyl phthalate (a feeding study), the dosed male rats died very early, so the controls were also sacrificed. These 7 groups were not included by Haseman *et al.* (4) because they had not lived long enough to yield a reasonable estimate of lifetime cancer risk, which was their outcome of interest. However, the 7 groups were examined in the same manner as the other control groups in this historical control data base and can thus be added to the data base for considering tumor onset as a function of time.

Intermediate Lethality Method

The prediction of time to tumor data usually requires detailed knowledge concerning the probable cause of death for animals with the tumor. A review of the statistical methods which apply to animals for which the probable cause of death is known is given by Kalbfleisch *et al.* (7). Unfortunately, in the NTP historical control data base, cause of death information is not available. In addition, even if the information on cause of death was available, its usefulness would be limited (8).

It is possible to obtain an upper bound on tumor prevalence by assuming that tumors are incidental; *i.e.*, they do not affect an animal's chances of dying (9). For many of the tumor types discussed here, the assumption that they occur in an incidental context is approximately correct. However, for diseases such as leukemia and lymphoma, this assumption appears to be incorrect. Portier (10) developed a method for fitting parametric models to the distribution of unobservable tumor onset times without assuming that the tumor is incidental or rapidly lethal (observable). This method is not as sensitive to tumor related differences in mortality when estimating the distribution of tumor onset times as the usual methods which must assume that all tumors are incidental or lethal. In addition, only a few sacrifice times are required in order to be able to estimate the tumor onset distribution. This method assumes that, for animals with the tumor, the probability of dying does not depend upon the age of the tumor but only upon the age of the animal. This "Markov" assumption will be discussed later.

Received 12/10/85; revised 4/28/86; accepted 6/4/86.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] The abbreviation used is: NTP, National Toxicology Program.

Oct. 2018 EPA-HQ-2018-000065 ED_001487_00013660-00001

This method of estimation will be referred to as the intermediate lethality method.

There is no literature providing a test of goodness of fit for this model using carcinogenicity study data (*i.e.*, left and right censored data). However, graphical methods have been used (11) to assess the fit of tumor prevalence models to the data from the ED_{01} study (12). This somewhat *ad hoc* method involves the comparison of the observed and expected prevalence at death for each observed death time (this is similar to methods used in standard regression problems).

A second method for testing the fit of the model is the standard χ^2 goodness of fit statistic. (For the statisticians: The data arise from a multinomial sampling procedure and we are using a method to estimate model parameters which is similar to the modified minimum χ^2 method. Thus, an approximate χ^2 statistic can be derived by taking the sum of the ratio of the square of the difference between the observed counts and the expected counts in each cell to the expected counts (*i.e.*, $\sum(O - E)^2/E$); see Ref. 13, Chap. 30). The degrees of freedom for the test would be s-r-1 where s is the number of sacrifice times and r is the number of parameters in the tumor onset model. This χ^2 test and the graphical approach mentioned above are used to determine the adequacy of the fit of the estimated model to the observed data.

To model tumor onset, a two-parameter Weibull model ($r = 2$) is used. This model is given as

$$S(t \mid \alpha,\beta) = e^{-\alpha t^{\beta}} \qquad \alpha > 0,\ \beta > 0 \tag{A}$$

where t is the age of the animal in weeks (not the number of weeks on study which is typically age minus 6 weeks) and $S(t \mid \alpha,\beta)$ represents the probability of not having a tumor at time t. This model has been used in many analyses to describe tumor onset as a function of time (4, 7, 11, 14–16).

Summary Measures of Lethality

Let $0 < t_1 < t_2 < \ldots < t_l$ denote the set of distinct death times. Let d_i^- denote the number of animals dying from causes other than sacrifice (natural causes) at time t_i which are found to be tumor free at death. Similarly, let d_i^+ denote the number of animals dying naturally with a tumor. Let $1 - \phi_i^-$ denote the probability of surviving natural death at time t_i given the animal has escaped death to this time and given the tumor has not occurred. Let $1 - \phi_i^+$ denote the probability of surviving natural death at time t_i given the animal has escaped death to this time and given the tumor is present.

The mortality parameters, ϕ^- and ϕ^+, provide an excellent method for determining the contribution of the tumor to mortality in the population. A measure commonly used to summarize the effect of presence or absence of some disease (tumor) on death is the relative risk. For each time t_i, the relative risk would be given by ϕ_i^+/ϕ_i^-. One obvious summary measure would be a weighted average of the relative risks. However, at times where no animals die without the tumor, the estimate of ϕ^- is zero and the relative risk is undefined. To avoid this problem, we will use the following measure of lethality which we will refer to as the relative average risk:

$$R = \frac{\sum_{i=1}^{l} (d_i^- + d_i^+)\phi_i^+}{\sum_{i=1}^{l} (d_i^- + d_i^+)\phi_i^-}$$

(*i.e.*, we take the ratio of a weighted average of the ϕ^+ to a weighted average of the ϕ^-). The weight at time t_i is the number of animals dying naturally at this time. This measure gives little weight to the hazards of death for early times and more weight to the hazards of death at times when most animals die. This avoids the problem of a few early deaths elevating the average lethality.

Subject to statistical variability, $R = 1$ indicates that the tumor is incidental, *i.e.*, it does not increase or decrease an animal's chances of dying. When $R > 1$, the tumor increases an animal's chances of dying. For example, if $R = 2$, then, on average, an animal with the tumor is twice as likely to die in the next time interval as an animal without the tumor. When $R < 1$, the tumor decreases the animal's chances of dying.

Table 1 *Early termination studies added to the NTP control data base*

NTP Toxicology Registry No.	Chemical name	Wk in study				No. in study			
		Mice		Rats		Mice		Rats	
		F	M	F	M	F	M	F	M
213	Butyl benzyl phthalate	112	112	111	35	50	50	49	50
234	Allyl isothiocyanate	86	86	83	83	50	50	50	50
250	Benzyl acetate	61	61	48	47	50	50	50	50
251	Toluene diisocyanate	87	87	86	86	50	49	50	50
252	Geranyl acetate	45	43	46	45	50	50	50	50
257	Diglycidyl resorcinol ether	62	62	59	58	50	50	50	50
263	1,2-Dichloropropane	41	41	40	40	50	50	50	50

RESULTS

Survival. Let $F(t)$ denote the probability of surviving to time t. Several models have been proposed for estimating survival in carcinogenicity experiments (17). In what follows, the method of maximum likelihood is used to estimate parametric curves for survival in the NTP historical control data base.

Several of the standard survival models were fit to the survival data from the NTP data base. These included the Weibull model (Equation A), the Gompertz model, the lognormal model, the Makeham-Gompertz model, and the log-logistic model. None of these "usual" models adequately fit the observed survival data. Looking at the Kaplan-Meier survival curves (18), it is evident that in all four sex/species groups the curves are linear in time for early deaths. Therefore, a modified Weibull model was considered. The modification added a linear parameter to $F(t)$ resulting in the form

$$F(t) = e^{-\alpha t - \gamma t^{\beta}} \qquad \alpha \geq 0,\ \gamma \geq 0,\ \beta > 0 \tag{B}$$

For small values of t and $\beta > 1$, this model behaves in a linear fashion. For large values of t and $\beta > 1$, this model behaves like the two-parameter Weibull model mentioned above. In all four cases this model predicted the observed response exceptionally well. Using a simple χ^2 test of observed minus expected, all four predicted models yielded acceptable fits of the data with test P values in the range of 0.3 to 0.6. The model parameters appear in Table 2.

The animals used in the historical control data base were generally sacrificed if they lived to 109 weeks and the latest death (due to sacrifice) was at 116 weeks. Thus, the application of these models beyond 109 weeks would be speculative. However, the NTP has conducted a series of bioassays on asbestos using a large control group of Fischer rats which were allowed to live until only 10% of the animals remained alive, at which point the remaining animals were sacrificed. The model parameters given in Table 2 adequately fit the asbestos controls through 140 weeks with only minor differences at approximately 100 weeks. Thus it is reasonable to conclude that the model (Equation B) and its associated parameter estimates given in Table 2 are appropriate to describe the general survival in rats for up to 140 weeks.

Tumor Onset Models. The next stage in analyzing these data was to provide estimates of the probability of having a tumor at all times between 0 and 116 weeks. This was done using the methods of Portier (10) and the Weibull model (Equation A).

Table 2 *Modified Weibull model parameters for general survival*

Species	Sex	α	γ	β
Fischer 344 rat	F	$1.2372\,e-4$	$2.4785\,e-16$	7.3839
	M	$1.2381\,e-4$	$9.0162\,e-17$	7.6671
$B6C3F_1$ mice	F	$2.8678\,e-4$	$1.4206\,e-14$	6.4978
	M	$6.5797\,e-4$	$1.2079\,e-15$	6.9593

Table 3 *Weibull model parameters for 2836 female B6C3F₁ mice*

Site (no. with tumor)	Model parameters[a] (SD) β	α	Goodness of fit χ^2 P
Circulatory system: hemangioma, hemangiosarcoma (77)	4.96 (6.45 e − 1)	4.49 e − 2 (6.46 e − 3)	<0.01
Digestive system: liver			
Adenoma (101)	2.55 (3.64 e − 0)	4.94 e − 2 (1.65 e − 2)	0.52
Carcinoma (104)	4.61 (1.48 e − 0)	5.95 e − 2 (8.71 e − 3)	0.38
Both (202)	3.39 (1.72 e − 0)	1.08 e − 1 (1.67 e − 2)	0.85
Endocrine system			
Pituitary all tumors (174)	12.21 (3.20 e − 0)	1.89 e − 1 (3.53 e − 2)	<0.01
Adrenal			
All pheochromocytoma (16)	11.64 (6.93 e − 0)	1.46 e − 2 (7.72 e − 3)	>0.99
Thyroid follicular cell			
Adenoma (41)	3.07 (6.75 e − 0)	2.31 e − 2 (1.49 e − 2)	<0.01
Adenoma and carcinoma (47)	3.53 (6.56 e − 0)	2.74 e − 2 (1.58 e − 2)	<0.01
Hematopoietic system: leukemia/lymphoma (699)	6.41 (5.16 e − 1)	5.01 e − 1 (2.44 e − 2)	0.04
Integumentary system: (neuro)fibroma and (neuro)fibrosarcoma (23)	4.93 (1.16 e − 0)	1.32 e − 2 (2.99 e − 3)	>0.99
Reproductive system: uterus			
Endometrial stromal polyp (23)	7.49 (1.65 e − 0)	1.63 e − 2 (3.72 e − 3)	>0.99
Respiratory system: lung			
Alveolar/bronchiolar adenoma (138)	2.93 (1.64 e − 0)	7.29 e − 2 (1.13 e − 2)	0.92
Alveolar/bronchiolar carcinoma (51)	3.58 (1.32 e − 0)	2.75 e − 2 (5.85 e − 3)	0.20
Both (188)	3.05 (1.38 e − 0)	1.01 e − 1 (1.25 e − 2)	0.90
All tumor bearing animals (1432)	4.56 (5.96 e − 1)	1.11 e − 0 (5.37 e − 2)	0.02

[a] Time is scaled by using age/118 weeks.

Table 4 *Weibull model parameters for 2692 male B6C3F₁ mice*

Site (no. with tumor)	Model parameters[a] (SD) β	α	Goodness of fit χ^2 P
Circulatory system: hemangioma, hemangiosarcoma (80)	9.22 (1.44 e − 0)	6.45 e − 2 (8.49 e − 3)	>0.99
Digestive system: liver			
Adenoma (261)	1.60 (4.90 e − 1)	1.29 e − 1 (9.43 e − 3)	0.91
Carcinoma (520)	3.79 (3.50 e − 1)	3.26 e − 1 (1.75 e − 2)	0.02
Both (766)	2.93 (3.95 e − 1)	4.83 e − 1 (2.25 e − 2)	0.41
Endocrine system			
Pituitary all tumors (12)	6.97 (1.70 e + 1)	9.99 e − 3 (2.04 e − 2)	0.72
Adrenal			
Cortical adenoma (53)	3.15 (7.15 e − 0)	2.99 e − 2 (2.30 e − 2)	0.60
Cortical adenoma and carcinoma (56)	3.18 (5.74 e − 0)	3.17 e − 2 (1.83 e − 2)	0.59
All pheochromocytoma (31)	1.47 (6.62 e − 1)	1.50 e − 2 (3.37 e − 3)	<0.01
Thyroid follicular cell			
Adenoma (26)	0.36 (2.11 e − 0)	1.11 e − 2 (2.84 e − 3)	0.77
Adenoma and carcinoma (32)	0.75 (2.06 e − 0)	1.46 e − 2 (4.46 e − 3)	0.76
Hematopoietic system: leukemia/lymphoma (312)	5.44 (9.22 e − 1)	2.10 e − 1 (1.74 e − 2)	0.72
Integumentary system: (neuro)fibroma and (neuro)fibrosarcoma (95)	7.95 (1.06 e − 0)	7.11 e − 2 (8.62 e − 3)	0.95
Respiratory system: lung			
Alveolar/bronchiolar adenoma (308)	2.07 (7.10 e − 1)	1.62 e − 1 (1.47 e − 2)	0.15
Alveolar/bronchiolar carcinoma (125)	5.67 (1.57 e − 0)	8.24 e − 2 (1.23 e − 2)	0.52
Both (425)	2.63 (8.80 e − 1)	2.41 e − 1 (2.06 e − 2)	0.04
All tumor bearing animals (1534)	3.42 (2.79 e − 1)	1.29 e − 0 (4.80 e − 2)	0.16

[a] Time is scaled by using age/118 weeks.

Tables 3 through 6 give the parameter estimates (and their standard deviations) for the two-parameter Weibull model and the P value for the χ^2 goodness of fit test. The standard deviations were calculated using the bootstrap estimates of variance (10) with a resample size of 100.

The tumors shown in Tables 3 to 6 are similar to those used by Haseman *et al.* (4). In some cases tumor types have been grouped together to provide sufficient tumor-bearing animals to use the intermediate lethality method. In addition certain tumors considered by Haseman *et al.* (4) are not listed since there were insufficient numbers of animals with these tumors to estimate tumor onset as a function of time for any of the four groups. The data were considered insufficient for fitting a parametric model using the intermediate lethality method if less than 10 animals had the tumor under study, or the first tumor was observed after 104 weeks, or less than 20 animals had the tumor *and* the first tumor was observed after 94 weeks. In all three of these cases, the quantal response is probably as good as most other measures in estimating the lifetime cancer risk in these animals. It is suggested the reader refer to Ref. 4 for the full set of tumors studied and the associated quantal response.

Table 5 *Weibull model parameters for 2719 female F344 rats*

Site (no. with tumor)	Model parameters[a] (SD) β	α	Goodness of fit χ^2 P
Digestive system: liver			
Neoplastic nodules (72)	14.34 (5.01 *e* − 0)	6.38 *e* − 2 (1.66 *e* − 2)	0.94
Neoplastic nodules and carcinoma (75)	14.71 (4.93 *e* − 0)	6.75 *e* − 2 (1.66 *e* − 2)	0.88
Endocrine system: pituitary			
All tumors (1114)	4.87 (5.69 *e* − 1)	8.79 *e* − 1 (3.87 *e* − 2)	0.52
Adrenal			
Cortical adenoma (76)	5.11 (4.96 *e* − 0)	4.42 *e* − 2 (1.40 *e* − 2)	0.08
Cortical adenoma and carcinoma (84)	5.07 (3.03 *e* − 0)	4.88 *e* − 2 (1.02 *e* − 2)	0.12
All pheochromocytoma (96)	3.69 (2.98 *e* − 0)	5.16 *e* − 2 (1.07 *e* − 2)	0.96
Thyroid			
C-cell adenoma (118)	2.59 (9.41 *e* − 1)	6.10 *e* − 2 (7.41 *e* − 3)	0.03
C-cell carcinoma (85)	4.21 (2.99 *e* − 0)	4.84 *e* − 2 (1.01 *e* − 2)	0.92
Both (200)	3.06 (7.48 *e* − 1)	1.09 *e* − 1 (9.98 *e* − 3)	0.26
Pancreatic islets			
All tumors (24)	9.24 (4.91 *e* − 0)	1.71 *e* − 2 (5.38 *e* − 3)	0.76
Hematopoietic system: Leukemia/lymphoma (461)	2.70 (4.26 *e* − 1)	2.44 *e* − 1 (1.24 *e* − 2)	0.29
Integumentary system: (neuro)fibroma and (neuro)fibrosarcoma (42)	7.19 (1.71 *e* − 0)	2.68 *e* − 2 (4.82 *e* − 3)	>0.99
Reproductive system: mammary gland fibroadenoma (594)	5.44 (1.00 *e* − 0)	3.94 *e* − 1 (2.65 *e* − 2)	0.60
Uterus: endometrial stromal polyp (463)	1.64 (3.59 *e* − 1)	2.32 *e* − 1 (1.20 *e* − 2)	0.01
Respiratory system: lung			
Alveolar/bronchiolar adenoma and carcinoma (31)	0.75 (2.71 *e* − 0)	1.28 *e* − 2 (3.90 *e* − 3)	<0.01
All tumor bearing animals (2074)	3.64 (2.49 *e* − 1)	2.14 *e* − 0 (5.99 *e* − 2)	0.31

[a] Time is scaled by using age/116 weeks.

Table 6 *Weibull model parameters for 2670 male F344 rats*

Site (no. with tumor)	Model parameters[a] (SD) β	α	Goodness of fit χ^2 P
Circulatory system: hemangioma, hemangiosarcoma (13)	5.11 (1.41 *e* − 0)	7.98 *e* − 3 (2.19 *e* − 3)	>0.99
Digestive system: liver			
Neoplastic nodules (79)	5.58 (2.70 *e* − 0)	5.13 *e* − 2 (1.15 *e* − 2)	0.80
Hepatocellular carcinoma (18)	4.91 (1.33 *e* − 0)	1.12 *e* − 2 (2.79 *e* − 3)	0.97
Both (97)	5.49 (2.24 *e* − 0)	6.29 *e* − 2 (1.20 *e* − 2)	0.81
Endocrine system: pituitary			
All tumors (535)	4.17 (7.17 *e* − 1)	3.74 *e* − 1 (2.44 *e* − 2)	0.01
Adrenal			
Cortical adenoma (28)	4.27 (7.09 *e* − 0)	1.66 *e* − 2 (1.07 *e* − 2)	0.95
Cortical adenoma and carcinoma (33)	4.48 (4.58 *e* − 0)	1.98 *e* − 2 (6.70 *e* − 3)	0.96
All pheochromocytoma (414)	5.55 (1.50 *e* − 0)	2.94 *e* − 1 (2.95 *e* − 2)	0.28
Thyroid			
C-cell adenoma (116)	7.85 (4.84 *e* − 0)	8.84 *e* − 2 (2.65 *e* − 2)	<0.01
C-cell carcinoma (88)	2.25 (1.04 *e* − 0)	4.65 *e* − 2 (6.35 *e* − 3)	0.06
Both (202)	3.70 (1.45 *e* − 0)	1.23 *e* − 1 (1.48 *e* − 2)	0.14
Follicular cell adenoma (22)	4.78 (1.83 *e* − 0)	1.34 *e* − 2 (3.31 *e* − 3)	0.70
Follicular cell carcinoma (17)	7.07 (4.13 *e* − 0)	1.22 *e* − 2 (4.24 *e* − 3)	>0.99
Both (39)	5.76 (1.62 *e* − 0)	2.59 *e* − 2 (4.53 *e* − 3)	0.83
Pancreatic islets: all tumors (131)	5.42 (2.60 *e* − 0)	8.74 *e* − 2 (1.60 *e* − 2)	0.56
Hematopoietic system: leukemia/lymphoma (702)	4.40 (7.33 *e* − 1)	4.76 *e* − 1 (2.69 *e* − 2)	0.10
Integumentary system: (neuro)fibroma and (neuro)fibrosarcoma (122)	3.85 (1.04 *e* − 0)	7.07 *e* − 2 (8.47 *e* − 3)	0.86
Reproductive system			
Mammary gland fibroadenoma (52)	7.14 (3.69 *e* − 0)	3.66 *e* − 2 (9.32 *e* − 3)	0.95
Testis interstitial cell (2099)	4.10 (1.28 *e* − 1)	3.51 *e* − 0 (1.18 *e* − 1)	<0.01
Respiratory system: lung			
Alveolar/bronchiolar adenoma (38)	1.00 (1.47 *e* − 0)	1.69 *e* − 2 (3.93 *e* − 3)	0.74
Carcinoma (21)	2.73 (3.21 *e* − 0)	1.10 *e* − 2 (3.75 *e* − 3)	0.91
Adenoma and carcinoma (57)	1.36 (1.56 *e* − 0)	2.65 *e* − 2 (5.00 *e* − 3)	0.72
Other tumors: mesothelioma, all types (54)	6.47 (1.14 *e* − 0)	3.65 *e* − 2 (5.78 *e* − 3)	0.98
All tumor bearing animals (exclude testis) (1191)	4.05 (4.08 *e* − 1)	8.94 *e* − 1 (3.36 *e* − 2)	<0.01

[a] Time is scaled by using age/116 weeks.

The Weibull model gave an adequate fit to the data for about 75% of the tumors studied. As mentioned before, the determination of goodness of fit was based upon a graphical analysis of the results and upon the χ^2 test. For every case in which the graphical analysis suggested a strong lack of fit, the χ^2 P was <0.05; therefore, it is sufficient to consider only the P values for the χ^2 test. It should be noted that in several cases, the P values for the χ^2 test are large (>0.95). In many of these cases, there were very few tumors (<25). The χ^2 test may not be valid for very small tumor rates and we caution the reader against overinterpretation of these fits. In the remaining cases where P is large, the graphical analysis supported the fact that the model seemed to agree very well with what was observed.

To understand the meaning of the values for the parameters in these tables, consider the following example. Suppose, prior to starting a study, that a researcher wishes to estimate the probability that untreated female $B6C3F_1$ mice will contract leukemia or lymphoma by 90 weeks of age. This is $1 - S(90/118 | \alpha = 0.501, \beta = 6.41)$ which equals 0.085 (the Weibull model parameters are given in Table 3). This means that after adjusting for survival differences, 8.5% of the animals would generally have the tumor by this time. If the tumor does not contribute to mortality in the animals ($R = 1$) then on average 8.5% of the untreated animals alive at 90 weeks should have the tumor. If the tumor is rapidly lethal ($R \gg 1$), about 8.5% of the animals placed on study should have died with the tumor by 90 weeks.

How can the researcher use this information once a study is begun? Since tumors can contribute to mortality without instantly killing the animal, determining whether the results from a given study are consistent with the model estimates from the historical control data will depend upon the degree to which the tumor contributes to mortality as well as the distribution of tumor onset times and its statistical variability. This is discussed in the next section.

In Tables 3 to 6, small values of β would indicate tumors which may occur early in untreated animals. In designing a study which is expected to yield an increase in tumors of this type, an early interim sacrifice may be informative in determining the tumor onset distribution as a function of time. Leukemia/lymphoma in female rats is such a tumor. For cases where β is large, the tumor is more likely to occur late in the animal's lifetime. For studies which expect to see an increase in tumors of this type, interim sacrifices may not be as useful for the purposes of detecting a tumor. Tumors of the pancreatic islets in female rats are of this type. Note that this is for treatment related increases in the parameter α from the Weibull model (Equation A) and not for treatment related changes in β. If it is expected that β will change as the treatment changes, interim sacrifices will be useful in most cases.

There are statistical factors which affect the use of these model parameters. The simultaneous estimation of α and β is difficult (14). Confidence regions for these parameters tend to be long and narrow allowing for a wide range of fits. The net effect of this problem in the present case is that even though the data set contains thousands of animals, in many cases a model with $\beta = 3$ is as likely to fit the data as a model with $\beta = 5$ (this is indicated by the large standard deviations for some tumors). Thus even though the model parameters presented in Tables 3 to 6 are the best estimates, it would be inappropriate to utilize these model parameters beyond the largest study time.

Tumor Lethality. Table 7 presents the relative average risks for the tumor sites presented in Tables 3 to 6. The number in parentheses is the standard deviation of the relative average risk based upon a bootstrap of 100 data sets. Large values for the relative average risk indicate (relatively) lethal tumors and values near or below 1 indicate nonlethal or incidental tumors.

Many of the values in Table 7 are near 1 suggesting that most tumors are incidental. Several tumors seem to be incidental in all species (*e.g.*, lung adenomas and carcinomas combined). Others, such as leukemia/lymphoma, seem lethal in all species. For female mice, there were four significantly lethal sites: circulatory system hemangiomas and hemangiosarcomas; leukemia/lymphoma; skin tumors [integumentary system (neuro)fibroma and (neuro)fibrosarcoma]; and all tumors combined [all tumor bearing animals]. With the exception of skin tumors, the lethality of these tumors is expected. Skin tumors provide a good example for examining the difference between biological lethality and statistical lethality. In general, a skin tumor may not directly cause the death of a mouse. However, the tumors can lead to weakening of the animals and infection. Because of this, these animals are more likely to be attacked by cage mates or to be sacrificed for humanitarian reasons (a moribund sacrifice). Thus, even though the tumor itself is not lethal, the presence of the tumor can lead to life shortening.

Pituitary tumors, adrenal pheochromocytomas, and lung carcinomas appear to be lethal in female mice, but their relative average risk is not statistically different from 1. All other tumors in female mice had an approximate relative average risk of 1.

Two tumors for male mice were different from those for female mice. Pituitary tumors did not appear to be lethal in male mice and liver carcinomas did appear to be lethal. For female rats, the results agreed with those for female mice with the exception of liver tumors, which were lethal, but not significantly. In male rats, liver tumors did not appear to be lethal and C-cell carcinomas of the thyroid and all mesotheliomas did appear to be lethal.

The estimates of lethality presented in Table 7 can be used to direct the analysis of results from a carcinogenicity study. If treatment does not affect the lethality of the tumor, then tumors which appear lethal or nonlethal in the controls will act in the same way in the treated groups. After studying the results in Table 7, one would feel comfortable with doing an incidental tumor analysis on mammary gland fibroadenomas in male and female Fischer rats; a life table analysis for this tumor would probably be inappropriate. On the other hand, a life table analysis would probably be more appropriate for integumentary system (neuro)fibromas and (neuro)fibrosarcomas than would be an incidental tumor analysis.

In comparing Tables 3 to 6 with Table 7, it is seen that for many of the more lethal findings mentioned above, the model did not fit the data. These very cases are where one would expect the model to fail. In cases where the tumor is nonlethal, the Markov assumption is correct and the goodness of fit test is comparing the Weibull model to the data. When lethality exists, the Markov assumption becomes critical since the correctness of both the Markov assumption and the Weibull model is being assessed. If the assumption is incorrect, the model will not fit the data and the magnitude of the relative average risk may be wrong, but the finding of lethality will probably not be wrong. We are currently evaluating the sensitivity of this method to deviations from this Markov assumption.

Now, by considering both tumor onset and lethality, it is possible to detect when the data diverge from the predicted model. Using the formulae provided by Portier (10) and the data it is possible to estimate all the ϕ_i from the estimates of α and β. It is then simple to estimate the expected prevalence which should equal the proportion of sacrificed animals with

Table 7 *Estimates of relative average risk (SD) from the NTP control data base*

Site	Female mice	Male mice	Female rats	Male rats
Circulatory system: hemangioma-hemangiosarcoma	2.79 (0.67)	12.19 (3.48)		4.03 (1.28)
Digestive system: liver				
Adenoma/neoplastic nodule	0.57 (0.80)	0.26 (0.08)	4.54 (3.03)	1.02 (0.42)
Carcinoma	1.29 (0.53)	2.90 (0.36)		1.62 (0.80)
Both	0.85 (0.38)	1.62 (0.19)	4.49 (2.62)	1.14 (0.40)
Endocrine system				
Pituitary: all tumors	4.29 (3.27)	1.01 (0.70)	1.80 (0.23)	1.66 (0.23)
Adrenal				
Cortical adenoma		1.12 (0.91)	0.80 (1.35)	1.23 (1.38)
Cortical adenoma and carcinoma		1.21 (0.81)	1.26 (1.10)	1.52 (0.91)
All pheochromocytoma	2.45 (1.29)	1.87 (0.75)	1.22 (0.98)	1.43 (0.35)
Thyroid				
C-cell adenoma			0.59 (0.20)	1.34 (1.67)
C-cell carcinoma			0.81 (0.38)	1.02 (0.26)
Both			0.68 (0.15)	3.70 (1.45)
Follicular cell adenoma	0.41 (1.07)		0.46 (0.67)	1.74 (0.84)
Follicular cell carcinoma				0.84 (0.78)
Both	0.37 (1.12)	1.03 (0.89)		1.35 (0.56)
Pancreatic islets: all tumors			3.04 (1.51)	1.46 (1.03)
Hematopoietic system: leukemia/lymphoma	6.26 (1.04)	4.52 (0.89)	5.09 (0.62)	5.05 (0.65)
Integumentary system: (neuro)fibroma and (neuro)fibrosarcoma	8.71 (2.06)	12.62 (3.21)	5.13 (1.91)	1.30 (0.30)
Reproductive system				
Mammary gland fibroadenoma			1.70 (0.31)	1.02 (0.76)
Uterus: endometrial stromal polyp	1.05 (0.74)		0.86 (0.12)	
Testis interstitial cell				0.56 (0.08)
Respiratory system: lung				
Alveolar/bronchiolar adenoma	0.89 (0.62)	0.43 (0.11)		1.14 (0.43)
Alveolar/bronchiolar carcinoma	3.41 (2.72)	2.38 (2.15)		0.56 (0.69)
Both	1.30 (0.97)	0.68 (0.18)	0.43 (0.29)	0.88 (0.41)
Other tumors: mesothelioma, all types				2.79 (1.47)
All tumor bearing animals	4.84 (1.55)	4.19 (0.62)	7.50 (1.88)	4.35 (0.54)

the tumor. At each sacrifice time it is then possible to detect divergence from the historical control response. The data and formulae for this approach are available from the first author upon request.

Even without these formulae, there is an approximate method for detecting divergence from the historical control onset rate. It is possible to obtain a crude upper confidence bound on the probability of tumor onset by using confidence bounds on the model parameters. A crude 97.5% confidence bound on a parameter can be obtained by adding or subtracting 2 SD from the estimate. An approximate upper bound on the probability of tumor onset is then obtained by substituting the lower bound for β and the upper bound for α into the survival function. For the leukemia/lymphoma example given above, the 97.5% upper bound on the probability of tumor onset at 90 weeks is $1 - S(90/118 \mid \alpha = 0.55, \beta = 5.38) = 0.12$. Now, using standard statistical techniques for binomial sampling we can define a critical region beyond which response is significant compared to the historical controls. For example, suppose this researcher enters 50 animals into study. If more than 12 of these 50 animals have been observed to have leukemia or lymphoma by 90 weeks the lower 97.5% confidence bound on the probability of tumor onset before 90 weeks would be greater than 0.12. This would indicate a significant increase in leukemia/lymphoma in this group as compared to the historical rate. Similar methods could be used to obtain lower bounds on the probability of tumor onset. Note that this should be done only to indicate a change over historical rates and that in virtually every case, the concurrent control group should be used for final inference.

DISCUSSION

Understanding the distribution of tumor onset times in untreated animals is critical to the analysis of lifetime rodent tumorigenicity studies. This paper attempts to give unbiased parametric estimates of what that distribution looks like for the rodent groups typically used by the National Toxicology Program. The parameter estimates given in this report are based upon a set of assumptions, most critical of which are the Markov assumption described in the text and the Weibull model (Equation A). As with any statistical analysis, if the assumptions are wrong, then it is possible all of the results are wrong including the parameter estimates and their standard deviations. It is impossible, using the data at hand, to determine if these assumptions are correct. However, we can try other assumptions and see if the results change. If the results do not change, the assumptions may have little effect on the eventual outcome (statisticians refer to this as a robust estimation procedure).

In an attempt to assess the sensitivity of the Weibull model parameters to the Markov assumption an alternative assumption was used to model the data. In this case it was assumed that the lethality of the tumor did not depend upon the age of the animal but only upon the age of the tumor. In many cases, the estimates obtained using this "semi-Markov" assumption gave results which were identical to those obtained using the Markov assumption. When the estimates differed, the estimates obtained using the Markov assumption always fit the data better than those obtained using the semi-Markov assumption. We are currently studying the possibility of using an assumption which combines the Markov and semi-Markov assumptions in

which the lethality of the tumor depends upon both the age of the tumor and the age of the animal. There is also empirical evidence that this assumption can be dropped altogether. This is being studied. Finally, it should be stressed that these estimates are valid only in the range of the data, from approximately 10 weeks of age to 118 weeks of age.

Since the distribution of onset times for most tumors could be adequately described using the two-parameter Weibull model, we did not fit the modified Weibull model (Equation 2) to the tumor onset data. We did fit the tumor onset data using a three-parameter Weibull model where instead of t^β we used $(t - \omega)^\beta$ for $\omega \geq 0$. Except for rare tumors, the results did not differ markedly for this model. For rare tumors, the addition of ω resulted in smaller estimates of β.

In summary, we have presented a set of models which accurately describe the survival in control Fischer 344 rats and control $B6C3F_1$ mice used by the NTP. These models have been shown to be accurate to 10% survival in the rats and to 116 weeks in the mice. In addition, for specific tumors, model parameters for the distribution of tumor onset times have been given with a discussion of the adequacy of the model for describing that tumor in that sex/species. Finally, for these tumors, we present statistical evidence concerning the relative risk of dying with the tumor.

REFERENCES

1. Hoel, D. Animal experimentation and its relevance to man. Environ. Health Perspect., *32:* 25–30, 1979.
2. Doll, R. The age distribution of cancer: implication for models of carcinogenesis. J. R. Stat. Soc. Ser. A, *134:* 133–166, 1971.
3. Whittemore, A., and Altshuler, B. Lung cancer incidence in cigarette smokers: further analysis of Doll and Hill's data for British doctors. Biometrics, *32:* 805–816, 1976.
4. Haseman, J., Huff, J., and Boorman, G. Use of historical control data in carcinogenicity studies in rodents. Toxicol. Pathol., *12:* 126–135, 1984.
5. Krewski, D., Crump, K., Farmer, J., Gaylor, D., Howe, R., Portier, C., Salsburg, D., Sielken, R., and Van Ryzin, J. A comparison of statistical methods for low dose extrapolation utilizing time-to-tumor data. Fundam. Appl. Toxicol., *3:* 140–160, 1983.
6. Dinse, G. Testing for a trend in tumor prevalence rates: I. Nonlethal tumors. Biometrics, *41:* 751–770, 1985.
7. Kalbfleisch, J., Krewski, D., and Van Ryzin, J. Dose response models for time to response toxicity data. Can. J. Stat., *11:* 25–50, 1983.
8. Lagakos, S., and Ryan, L. On the representativeness assumption in prevalence tests of carcinogenicity. Appl. Stat., *34:* 54–62, 1985.
9. Peto, R., Pike, M., Day, N., Gray, R., Lee, P., Parish, S., Peto, J., Richards, S., and Wahrendorf, J. Guidelines for simple sensitive significance tests for carcinogenic effects in long term animal experiments. *In*: Long-Term and Short-Term Screening Assays For Carcinogens: A Critical Appraisal, IARC Monographs, Annex to Supplement 2, pp. 311–426, Lyon, France: International Agency for Research on Cancer, 1980.
10. Portier, C. Estimating the tumor onset distribution in animal carcinogenesis experiments. Biometrika, in press, 1985.
11. Brown, K., and Hoel, D. Modeling time-to-tumor data: analysis of the ED_{01} study. Fundam. Appl. Toxicol., *3:* 458–469, 1983.
12. Staffa, J., and Mehlman, M. Innovations in Cancer Risk Assessment, Park Forest South, Il: Pathotox Publishers, 1979.
13. Cramer, H. Mathematical Methods of Statistics. Princeton, NJ: Princeton University Press, 1971.
14. Peto, R., and Lee, P. Weibull distributions for continuous carcinogenesis experiments. Biometrics, 29: 457–470, 1973.
15. Portier, C., Hoel, D., and Van Ryzin, J. Statistical analysis of the carcinogenesis bioassay data relating to the risks from exposure to 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. *In*: W. Lowrance (ed.), Public Health Risks of the Dioxins. Los Altos, CA: William Kaufmann, Inc., 1984.
16. Peto, R., Gray, R., Brantom, P., and Grasso, P. Nitrosamine carcinogenesis in 5120 rodents. IARC Meeting on Nitrosamines, Banff, Canada, September 4–9, 1983. Lyon, France: International Agency for Research on Cancer, 1984.
17. Elandt-Johnson, R., and Johnson, N. Survival Models and Data Analysis. New York: John Wiley & Sons, 1980.
18. Kaplan, E., and Meier, P. Nonparametric estimation from incomplete observations. J. Am. Stat. Assoc., 53: 457–481, 1958.