

Estimation of Lag Time Between Onset of and Death from an Occult Tumor

Hongshik Ahn¹, Hojin Moon² and Ralph L. Kodell³

¹Department of Applied Mathematics and Statistics
Stony Brook University, Stony Brook, NY 11794-3600

²Department of Mathematics and Statistics
California State University, Long Beach, CA 90840-1001

³Department of Biostatistics
University of Arkansas for Medical Sciences, Little Rock, AR 72205

Running title: Estimation of Lag Time from an Occult Tumor

***Corresponding author:**

Hongshik Ahn, Ph.D.
Department of Applied Mathematics and Statistics
Stony Brook University
Stony Brook, NY 11794-3600
Tel: 631-632-8372
Email: hahn@ams.sunysb.edu

Abstract

A new statistical method for estimating the lag time between onset of and death from an occult tumor is proposed for data without cause-of-death information. In this method, the survival function for time to tumor onset, tumor-specific survival function, and competing risks survival function are estimated using the maximum likelihood estimates of the parameters. The proposed method utilizes the estimated survival functions and statistically imputed fatal tumors to estimate the lag time. This approach is developed for rodent tumorigenicity assays that have at least one interim sacrifice and a terminal sacrifice. If the data contains cause-of-death information given by pathologists and it is believed to be reliable, it may be used for estimating the lag time. The proposed method is illustrated using a real data set. The accuracy of the estimation of lag time is evaluated via a Monte Carlo simulation study. Our study shows that the estimated lag time is quite reliable.

KEY WORDS: Bioassay, Cause of death, Likelihood, Sacrifice, Tumor lethality

1. Introduction

Animal carcinogenicity experiments are employed to test the carcinogenic potential of drugs and other chemical substances used by humans. Many methods have been proposed for analyzing tumor incidence data from animal bioassays. Kodell and Nelson (1980) presented a parametric model to estimate the average time to death from a specific tumor after onset for data with cause-of-death information and interim sacrifices. However, the elapsed-time (lag-time) estimation has not been done extensively by other

researchers. In this paper, a method is proposed for finding a representative estimate of the lag time in conjunction with the estimates of the survival functions and imputed number of fatal tumors proposed by Ahn, Kodell and Moon (2000) and modified by Moon, Ahn and Kodell (2002). The proposed method for lag-time estimation is based on a nonparametric model, and can be applied to data with either interim sacrifices. It is designed for data lacking cause-of-death information, although it can utilize the cause-of-death assigned by pathologists.

A survival/sacrifice experiment with benzidine dihydrochloride was conducted at the National Center for Toxicological Research to study strain and sex differences with regard to chemically-induced liver tumors in mice. The data were previously introduced and analyzed by Kodell and Nelson (1980). Male and female mice from two strains (F1 and F2) were used. The data from dose groups 60 ppm, 120 ppm, 200 ppm and 400 ppm were reported in Kodell and Nelson (1980). The counts of sacrifices and deaths with and without the tumor of interest are given in Table 1. Figure 1 shows the difference between the time-to-tumor-onset survival function $S(t)$ and the tumor-specific survival function $P(t)$ for each strain and sex. These survival functions will be discussed in Section 2.

The lag-time estimation requires cause-of-death information. Currently, cause-of-death information is obtained from case-by-case assignment of context of observation by pathologists in animal carcinogenicity experiments. In practice, however, accurate determinations of the cause of death are not easy, and classification errors can produce biases (Lagakos, 1982; Racine-Poon and Hoel, 1984; Lagakos and Louis, 1988). In this study, the imputed cause of death by Moon et al. (2002) is used as a replacement of the cause-of-death assignment by pathologists.

In estimating lag time, the control and dose groups are pooled for estimating the parameters, since it is desirable to estimate one representative lag time for a specific tumor. We expect to obtain a better lag-time estimate by pooling the data by using more information. The analysis by Kodell and Nelson (1980) showed that the average tumor onset time varies for the different dose groups, but the average time to death from a specific tumor after onset does not change substantially for different dose groups.

In this paper, the maximum likelihood estimates of the parameters obtained by Kim et al. (2007) and the imputed number of fatal and incidental tumors obtained by Ahn et al. (2000) are used in the proposed method for estimating the lag time. Alternatively, the number of fatal and incidental tumors assigned by pathologists can be used for the lag-time estimation, although we recommend using the statistically imputed numbers.

2. Estimation of Time to Onset

For the purpose of estimating the lag time, combine the data in control and all the dose groups. Divide the time scale into discrete intervals, with the j th interval given by $(t_{j-1}, t_j]$, $j = 1, \dots, J$, where $t_0 = 0$ and t_j denote the terminal sacrifice time. In the case of an experiment with multiple sacrifices, sequential sacrifice is performed at the end of each interval. For data with a single terminal sacrifice, pre-determined intervals such as the NTP (National Toxicology Program) intervals (Bailer and Portier, 1988) can be imposed.

Using the notation of Moon et al. (2002), define $S(t)$ as the survival function for the distribution of time to tumor onset, i.e., $S(t) = \Pr(T_1 > t)$, where T_1 is the random variable representing time to onset of the tumor of interest. Time to onset is defined as

the time when an occult tumor first becomes large enough to be detected histologically. Define $P(t)$ as the survival function with respect to death caused by the tumor of interest, i.e., $P(t) = \Pr(T_D > t)$, where T_D represents the overall time to death from the tumor of interest. Define $Q(t)$ as the survival function with respect to death from competing risks, i.e., $Q(t) = \Pr(X_C > t)$, where X_C is the random variable representing time to death from a cause other than the tumor of interest. Moon et al. (2002) formulated the contributions to the likelihood function in terms of $S(t)$, $P(t)$ and $Q(t)$ for multiple-sacrifice data as shown in Table 2, as if complete information on cause of death is available for the experimental animals. Following Malani and Van Ryzin (1988) (see also Kodell and Ahn, 1997), they assumed that T_1 and T_D occur before X_C in an interval and that probabilities involving T_1 , T_D and X_C can only be evaluated at the end of intervals. For experiments with a single terminal sacrifice, a_{2j} and b_{2j} are zero except for the last interval. Further details on these likelihood contributions are given in Moon et al (2002).

For all j , let $\pi_j = S(t_j)/P(t_j)$, $p_j = P(t_j)/P(t_{j-1})$, and $q_j = Q(t_j)/Q(t_{j-1})$. Then the survival functions can be expressed as

$$P(t_j) = \prod_{k=1}^j p_k, \quad Q(t_j) = \prod_{k=1}^j q_k, \quad \text{and} \quad S(t_j) = \pi_j P(t_j), \quad j=1, \dots, J. \quad (1)$$

After the reparameterization, Kim et al. (2007) obtained the log-likelihood function as

$$\begin{aligned} l = \sum_{i=1}^J & [\{N_j(\log p_j + \log q_j) + (d_j + a_{1j})\log\{1 - p_j + (1 - q_j)p_j(1 - \pi_j)\} \\ & + b_{1j}\{\log p_j + \log(1 - q_j) + \log \pi_j\} + a_{2j}\{\log p_j + \log q_j + \log(1 - \pi_j)\} \\ & + b_{2j}\{\log p_j + \log q_j + \log \pi_j\}] + \text{constant}, \end{aligned}$$

where N_j is the number of live animals at t_j . The monotonicity of $S(t)$, $P(t)$ and $Q(t)$ provides the following inequality conditions:

$$0 \leq \pi_j \leq 1, \quad 0 \leq p_j \leq 1, \quad 0 \leq q_j \leq 1, \quad \text{and } \pi_j p_j \leq \pi_{j-1} \text{ for } j = 1, \dots, J. \quad (2)$$

Kim et al. (2007) analytically derived the constrained MLEs of π_j , p_j , and q_j , as follows:

Case 1

$$\text{If } \frac{b_{2j}}{a_{2j} + b_{2j}} \leq \frac{b_{1j}}{d_j + a_{1j} + b_{1j}}, \text{ then the MLEs are } \begin{cases} \hat{\pi}_j = \max \left\{ \frac{b_{1j} + b_{2j}}{N_{j-1} - N_j}, \hat{\pi}_{j+1} \hat{p}_{j+1} \right\} \\ \hat{p}_j = 1 \\ \hat{q}_j = \frac{N_j^*}{N_{j-1}} \end{cases}$$

Case 2

$$\text{If } \frac{b_{1j}}{d_j + a_{1j} + b_{1j}} \leq \frac{b_{2j}}{a_{2j} + b_{2j}}, \text{ then the MLEs are } \begin{cases} \hat{\pi}_j = \max \left\{ \frac{b_{2j}}{a_{2j} + b_{2j}}, \hat{\pi}_{j+1} \hat{p}_{j+1} \right\} \\ \hat{q}_j = \frac{N_j^* \hat{\pi}_j}{N_j^* \hat{\pi}_j + b_{1j}} \\ \hat{p}_j = \frac{N_j^* + b_{1j}}{N_{j-1} \{ \hat{q}_j + (1 - \hat{q}_j) \hat{\pi}_j \}} \end{cases}$$

Here, we define N_j^* as the number of live animals right before t_j . Thus $N_j^* = N_j + a_{2j} + b_{2j}$ and $N_{j-1} = d_j + a_{1j} + b_{1j} + N_j^*$. Further details are given in Kim et al. (2007).

3. Estimation of Lag Time

In this section, a new method is developed for calculating a representative estimate of the elapsed time (lag time) between the time to onset of an occult tumor and the time that it causes the death of an animal.

To estimate the lag time, several notations will be introduced. Let Tb_j be the number of tumor bearing animals at the j th interval; then Tb_j can be expressed as $a_{1j} + d_j + a_{2j} + N_j Q(t_j) \{P(t_j) - S(t_j)\}$. Let To_j be the number of tumors that occurred at the j th interval, and let Te_j be the number of deaths or sacrifices with tumor at the j th interval. Define $lag_j(k)$ as the number of deaths at the j th interval from the tumor of interest having interval lag k (tumor occurred at the $(j-k)$ th interval), $k = 0, 1, \dots, j-1$. For the first interval, it is clear that To_1 is the same as Tb_1 and for $j \geq 2$,

$$Tb_j = To_j + Tb_{j-1} - Te_{j-1} = To_j + N_{j-1} Q(t_{j-1}) \{P(t_{j-1}) - S(t_{j-1})\}. \quad (3)$$

From Equation (3), To_j can be obtained as $Tb_j - N_{j-1} Q(t_{j-1}) \{P(t_{j-1}) - S(t_{j-1})\}$. To avoid having a negative estimate of To_j , Tb_j should be larger than $N_{j-1} Q(t_{j-1}) \{P(t_{j-1}) - S(t_{j-1})\}$. In order to make an adjustment of Tb_j and To_j , first define

$$A_j = \min\{N_j Q(t_j) [P(t_j) - S(t_j)], Tb_{j+1}\}, \quad j = 1, m-1.$$

For the last interval, find $Tb_m = a_{1m} + d_m + a_{2m} + N_m Q(t_m) \{P(t_m) - S(t_m)\}$ and calculate $To_m = Tb_m - A_{m-1}$. Starting from interval $m-1$, obtain Tb_j and then To_j as

$$Tb_j = a_{1j} + d_j + a_{2j} + A_j \text{ and } To_j = Tb_j - A_{j-1}.$$

For the first interval, obtain $Tb_1 = a_{11} + d_1 + a_{21} + A_1$ and $To_1 = Tb_1$. This approach guarantees nonnegative estimate of To_j , $j = 1, \dots, m$.

The estimator of $\text{lag}_j(k)$ is given as

$$\text{lag}_j(k) = d_j \prod_{r=j-k+1}^j \frac{Tb_r - To_r}{Tb_r} \frac{To_{j-k}}{Tb_{j-k}}$$

for $k = 0, 1, \dots, j-1$. The estimated number of death with fatal tumors from Section 2 can substitute for d_j in the above equation. Appendix A.1 provides further details of the estimation of lag_j .

Let u_j be the length of the j th interval, $j = 1, \dots, m$. If an animal died at the j th interval from the tumor of interest which occurred at the same interval, then the expected lag time of the tumor for that animal is

$$u_j / 4, \tag{4}$$

if we assume that both the time of tumor onset and the time of death from the tumor are distributed as a uniform in the interval. If $u_1 = \dots = u_m = u$, then the expected lag time is $u/4$. If an animal obtained the tumor at the r th interval and died from the tumor at the j th interval such that $j > r$, then the expected lag time of the tumor for that animal is

$$\frac{u_r + u_j}{2} + \sum_{s=r+1}^{j-1} u_s, \quad j = r+1, r+2, \dots, m. \tag{5}$$

If $u_1 = \dots = u_m = u$, then the expected lag time is $(j-r)u$. Further detail of the individual lag time is given in Appendix A.2.

Now we develop an estimation of the lag time between the time to onset of an occult tumor and the time that it causes the death of an animal. For the fatal tumors whose onset occurred at each interval, a method of estimating the average survival time

after onset is introduced. Using these estimates, the average lag time for the tumor, i.e., the average of $T_D - T_1$, is obtained.

Define tlag_r as the average survival time for the animals with fatal tumors whose onset occurred at the r th interval. If $r = m$, from (4), $\text{tlag}_m = u_m / 4$. For the fatal tumors whose onset occurred at the r th interval, the number of deaths is $\text{lag}_j(j - r)$ if the tumor caused death at the j th interval. If $r < m$, using (4) and (5), we obtain

$$\text{tlag}_r = \frac{u_r \text{lag}_r(0) / 4 + \sum_{j=r+1}^m [\{(u_r + u_j) / 2\} + \sum_{s=r+1}^{j-1} u_s] \text{lag}_j(j - r)}{\sum_{j=r}^m \text{lag}_j(j - r)}. \quad (6)$$

If $u_1 = \dots = u_m = u$, then the average survival time is

$$\text{tlag}_r = \frac{u \left\{ \text{lag}_r(0) / 4 + \sum_{j=r+1}^m (j - r) \text{lag}_j(j - r) \right\}}{\sum_{j=r}^m \text{lag}_j(j - r)}.$$

A representative estimate of the lag time between onset of and death from an occult tumor can be estimated by averaging the average survival times obtained above. Let L be the overall average lag time, then it can be estimated as

$$L = \sum_{r=1}^m w_r \text{tlag}_r, \quad (7)$$

where $w_r = u_r / \sum_{j=1}^m u_j$.

4. Example: Benzidine Data

The proposed methods are applied to analyze the Benzidine dihydrochloride data discussed in Section 1. The four dose groups are combined in order to estimate the lag time. However, the F1 strain Male 60 ppm dose group is not included due to insufficient

observations for obtaining estimates. The other three dose groups are combined for this one. The number of deaths from fatal tumor was estimated using the proposed method and compared with the numbers assigned by the pathologists. Frequency data and estimated quantities of interest along with the estimates of Tb_j and To_j are given in Table 3. The cause of death (COD) assigned by pathologists and the estimated COD are compared in this table.

Table 4 presents the estimates of the time-to-onset survival function, $S(t)$, from Table 3, along with estimates of a similar tumor onset survival function obtained from a discrete cumulative tumor incidence function produced by the method of Kodell and Ahn (1997). Note that the method of Kodell and Ahn does not involve cause of death. This table shows that our estimates of $S(t)$ are very close to that of Kodell and Ahn. A comparison of these estimates serves as a validity check on the method developed in this paper, because the method of Kodell and Ahn is completely nonparametric and makes fewer assumptions. The good correspondence between the two sets of estimates lends credibility to the estimates of the lethality and the tumor onset survival function, $S(t)$.

Table 5 compares the estimated lag time obtained using the proposed method and the method by Kodell and Nelson (1980). The estimated lag time by our method is approximately 4 to 6 months after onset. The estimated lag times from our method are quite close to those by Kodell and Nelson except for the F1 male group. Note that $d_1 = 1$ in the F1 male group, while $d_1 = 0$ in the other three groups according to both our method and the assignment by the pathologists. When we compare F1 and F2 male groups, there is a substantial difference in $tlag_1$, while $tlag_2$ and $tlag_3$ are close between

the two groups. This difference is caused by the death from the tumor occurred in the first interval of the F1 male group. In the F1 male group, $t\text{lag}_1$ is calculated as

$$t\text{lag}_1 = \frac{10(.8966) + 30(.1833) + 50(.4126)}{1.4925} = 23.5145$$

by substituting $\text{lag}_r(j)$ from Table 6 into Equation (6). In a hypothetical case that $\hat{d}_1 = 0$ instead of $\hat{d}_1 = 1$ in this group, the value of $t\text{lag}_1$ becomes

$$t\text{lag}_1 = \frac{30(.1833) + 50(.4126)}{.5959} = 43.8480$$

which is very close to the value of $t\text{lag}_1$ in the F2 male group. Further, $\hat{L} = (43.480/2) + (15.3858 + 5)/4 = 27.0$ becomes close to the estimate of the overall average lag time 26.3 in the F2 male group.

Our estimates appear to be reliable based on the results in Table 4 and the simulation results given in Section 5. The tumor observed in this example can be considered highly lethal according to both the assignment by pathologists and our estimation method. Our estimated lag time is close to the actual lag time according to our simulation study, especially for intermediate and highly lethal tumors.

5. Simulation

A Monte Carlo simulation study was conducted to evaluate the accuracy of the estimation of the lag time. A bioassay design with four dose groups (0, 1, 2, 4) of 50 animals each was considered. The proposed procedure was simulated to have the NTP intervals (0-52, 53-78, 79-92 and 93-104 weeks) with interim sacrifices at 52, 78 and 92 weeks and a terminal sacrifice at 104 weeks, which is the normal term of a chronic two-year study in rodents. The interim sacrifices were implemented by rescheduling a portion

of the 50 animals per group for early sacrifice, which are normally scheduled for a single terminal sacrifice in the customary lifetime rodent bioassay. Six animals were randomly pre-selected to be sacrificed at the end of each interval. All the remaining live animals were sacrificed at the end of the experiment.

It was assumed that three independent random variables completely determined the observed outcome for each animal. The random variables were the time to onset of tumor, T_1 , the time after onset until death from the tumor, T_2 , and the time to death from a competing risk, X_C . Note that $T_1 + T_2 = T_D$, where T_D represents the overall time to death from the tumor of interest. Thus the tumor of interest was present in an animal at the time of death if $T_1 \leq \min\{X_C, X_S\}$, where X_S denotes an animal's scheduled sacrifice time. An animal died from the tumor of interest if $T_D \leq \min\{X_C, X_S\}$. Otherwise, it died from a competing risk including sacrifice. For the animals die from the tumor of interest, the lag time is $T_2 = T_D - T_1$. Therefore, in this simulation, the estimated lag time of the tumor was compared with the average of the simulated values of T_2 for each simulation data set.

Distributions of time to onset, time to from the tumor of interest, and time to death from competing risks were of the form used by Moon et al. (2002). The probability of the background tumor onset probability (for the control group) by 104 weeks was chosen to be 0.05 (rare tumor), 0.15 or 0.3 (common tumor). The probability of tumor probability of tumor onset was chosen such that tumor onset probability in the highest dose group by 104 weeks was 5, 3 and 2 times the back ground tumor rates 0.05, 0.15 and 0.3, respectively. The tumor lethality rates were chosen to be approximately 5%, 35% and 60%.

Lag time was calculated after pooling the four dose groups. Five thousand simulated data sets with Weibull tumor onset distributions (with shape parameter 3) were generated with the competing risks survival rate 0.5 and two tumor lethality rates. Table 7 shows the simulation results. The average of the true lag times and the average of the estimated lag times over the 5000 simulation data sets are calculated for each configuration.

It is difficult to obtain an accurate estimation of lag time if the tumor is mostly incidental because of a lack of information on death caused from the tumor. However, our estimated lag time appears to have an agreement with the actual lag time in this table as long as the tumor is moderately or highly lethal.

6. Concluding Remarks

It is not an easy task to estimate the tumor onset time for an occult tumor. Kodell and Nelson (1980) developed a semi-Markov model for describing the development of a tumor which caused a death. Using this model, they provided the average time to tumor and time to death from tumor after onset. This method requires cause-of-death information. In many animal bioassays for carcinogenicity, however, cause-of-death information is not available. Even when cause of death is available, it is subject to error, potentially leading to misrepresentative estimates and tests of tumor rates. The present paper develops a method of estimating the lag time between the tumor onset and the death from that tumor, without the need for cause-of-death data. The proposed estimating method utilizes the distributions of time to tumor and time to death from the tumor in

conjunction with the distribution of competing risks given by Moon et al. (2002). The cause of death is statistically attributed by the method of Ahn et al. (2000).

Most of the existing testing procedures do not provide accompanying estimates of the tumor onset distribution function, while such an estimate is particularly useful for graphical displays to illustrate the results of the statistical tests. The proposed method provides the tumor onset distribution as well as the survival function for the tumor. The proposed lag time is estimated based on these survival functions and it is reasonably accurate especially when the tumor is moderately or highly lethal since the data contain enough information on animals died from the tumor.

Acknowledgments

Hongshik Ahn's research was partially supported by the Faculty Research Participation Program at the NCTR administered by the Oak Ridge Institute for Science and Education through an interagency agreement between USDOE and USFDA. Hojin Moon's research was partially supported by the Scholarly and Creative Activities Committee (SCAC) Award from CSULB.

References

- Ahn, H., Kodell, R. L., and Moon, H. (2000). Attribution of tumor lethality and estimation of time to onset of occult tumors in the absence of cause-of-death information. *Applied Statistics* 49:157-169.
- Bailer, A. J. and Portier, C. J. (1988). Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples. *Biometrics* 44:417-431.
- Kim, W., Ahn, H. and Moon, H. (2007). A dose-response test via closed-form solutions for constrained MLEs in survival/sacrifice experiments. *Statistics in Medicine* 26(3):694-708.
- Kodell, R. L. and Ahn, H. (1997). Age-adjusted trend test for the tumor incidence rate. *Biometrics* 53:1467-1474.
- Kodell, R. L. and Nelson, C. J. (1980). An illness-death model for the study of the carcinogenic process using survival/sacrifice data. *Biometrics* 36:267-277.
- Lagakos, S. W. (1982). An evaluation of some two-sample tests used to analyze animal carcinogenicity experiments. *Utilitas Mathematica* 21B:239-260.
- Lagakos, S. W. and Louis, T. A. (1988). Use of tumour lethality to interpret tumorigenicity experiments lacking cause-of-death data. *Applied Statistics* 37(2):169-179.
- Malani, H. M. and Van Ryzin, J. (1988). Comparison of two treatments in animal carcinogenicity experiments. *Journal of the American Statistical Association* 83: 1171-1177.

Moon, H., Ahn, H. and Kodell, R. L. (2002). Extension of Peto's test by attribution of tumor lethality in the absence of cause-of-death information. *Biometrical Journal* 44:982-1001.

Racine-Poon, A. and Hoel, D. G. (1984). Nonparametric estimation of the survival function when cause of death is uncertain. *Biometrics* 40:1151-1158.

Appendix

A. 1. Estimation of $\text{lag}_j(k)$

1. For $j = 1$, $\text{lag}_1(0) = d_1$, because $Tb_1 = To_1$.
2. For $j = 2$,
 - a. If $Tb_2 = To_2$ (i.e., $Tb_1 = Te_1$), then $\text{lag}_2(0) = d_2$ and $\text{lag}_2(1) = 0$.
 - b. If $Tb_2 > To_2$ (i.e., $Tb_1 > Te_1$), then among the tumor bearing animals in this interval, To_2/Tb_2 of them obtained tumor in the second interval and $(Tb_2 - To_2)/Tb_2$ of them obtained tumor in the first interval. Therefore,

$$\text{lag}_2(0) = d_2 \frac{To_2}{Tb_2},$$

$$\text{lag}_2(1) = d_2 \frac{Tb_2 - To_2}{Tb_2} = d_2 N_1 Q(t_1) \{P(t_1) - S(t_1)\} / Tb_2$$

from (3). Note that

$$\begin{aligned} \frac{To_j}{Tb_j} &= \Pr(\text{Tumor occurring at } j \mid \text{Tumor bearing at } j) \\ &= \Pr(j-1 < T_1 \leq j \mid T_1 \leq j, T_D > j-1, X_c > j-1), \\ \frac{Tb_j - To_j}{Tb_j} &= \Pr(\text{Tumor occurring before } j \mid \text{Tumor bearing at } j) \\ &= \Pr(T_1 \leq j-1 \mid T_1 \leq j, T_D > j-1, X_c > j-1). \end{aligned}$$

3. For $j = 3$,

a. If $Tb_3 = To_3$, then $\text{lag}_3(0) = d_3$, $\text{lag}_3(1) = 0$ and $\text{lag}_3(2) = 0$.

b. If $Tb_3 > To_3$,

i. If $Tb_2 = To_2$,

$$\text{lag}_3(0) = d_3 \frac{To_3}{Tb_3},$$

$$\text{lag}_3(1) = d_3 \frac{Tb_3 - To_3}{Tb_3},$$

$$\text{lag}_3(2) = 0.$$

ii. If $Tb_2 > To_2$,

$$\text{lag}_3(0) = d_3 \frac{To_3}{Tb_3},$$

$$\text{lag}_3(1) = d_3 \frac{Tb_3 - To_3}{Tb_3} \frac{To_2}{Tb_2},$$

$$\text{lag}_3(2) = d_3 \frac{Tb_3 - To_3}{Tb_3} \frac{Tb_2 - To_2}{Tb_2}.$$

c. In general, for the j th interval,

$$\text{lag}_j(0) = d_j \frac{To_j}{Tb_j},$$

$$\text{lag}_j(1) = d_j \frac{Tb_j - To_j}{Tb_j} \frac{To_{j-1}}{Tb_{j-1}},$$

$$\text{lag}_j(2) = d_j \frac{Tb_j - To_j}{Tb_j} \frac{Tb_{j-1} - To_{j-1}}{Tb_{j-1}} \frac{To_{j-2}}{Tb_{j-2}}.$$

$$\text{lag}_j(k) = d_j \prod_{r=j-k+1}^j \frac{Tb_r - T_{O_r}}{Tb_r} \frac{T_{O_{j-k}}}{Tb_{j-k}}, \quad k = 0, 1, \dots, j-1.$$

A. 2. Estimation of individual lag time

1. Suppose a tumor and death (from the tumor) occurred at the j th interval for an animal. Define $V = T_1 - \sum_{s=1}^{j-1} u_s$ and $W = T_D - \sum_{s=1}^{j-1} u_s$, then $W - V = T_D - T_1$ is the lag time of the tumor for that animal. We assume that tumors precede deaths, but the exact time of the tumor onset and death are not known. Thus we can assume V is distributed as $U(0, u_j)$ and given V , W is distributed as $U(V, u_j)$, where $U(a, b)$ denotes the uniform distribution on the interval (a, b) . The joint pdf of W and V is given as

$$f_{W,V}(w, v) = f_{W|V}(w|v) f_V(v) = 1/[u_j(u_j - v)], \quad 0 < v < w < u_j.$$

Let $X = W - V$, then by an appropriate transformation of the variables, we obtain the pdf of x given by

$$f(x) = -\frac{1}{u_j} \log \frac{x}{u_j}, \quad 0 < x < u_j,$$

and the expectation $E(X) = u_j / 4$.

2. Suppose a tumor occurred in an animal at the r th interval and the animal died from the tumor at the j th interval such that $j > r$. Define $V = T_1 - \sum_{s=1}^{r-1} u_s$ and $W = T_D - \sum_{s=1}^{r-1} u_s$, then $W - V = T_D - T_1$ is the lag time of the tumor for that animal. Assume that V is distributed as $U(0, u_r)$ and W is distributed as $U\left(\sum_{s=r}^{j-1} u_s, \sum_{s=r}^j u_s\right)$. The estimated lag time is

$$E(W - V) = E(W) - E(V) = \frac{u_r + u_j}{2} + \sum_{s=r+1}^{j-1} u_s, \quad j = r+1, \dots, m.$$

Table 1: Frequency data from Benzidine data

Strain	Group	j^1	Female				Male			
			Natural with ²	w/o ³	Sacrifice with	w/o	Natural with	w/o	Sacrifice with	w/o
F1	60ppm	1	0	0	1	71	* ⁴			
		2	0	1	3	45				
		3	6	4	18	18				
	120ppm	1	0	1	0	48	0	0	0	48
		2	4	1	14	31	2	5	7	37
		3	15	4	25	1	2	1	11	31
	200ppm	1	0	2	4	43	0	4	3	44
		2	6	0	38	7	2	2	5	37
		3	16	0	4	0	2	2	8	11
	400ppm	1	0	0	16	8	1	0	0	24
		2	29	0	9	1	0	2	11	11
		3	8	0	0	0	6	0	11	4
F2	60ppm	1	0	2	0	70	0	8	0	68
		2	0	1	10	38	0	3	3	41
		3	7	4	15	20	3	3	4	34
	120ppm	1	0	7	2	44	0	3	0	46
		2	6	4	15	26	1	1	8	39
		3	16	2	20	1	3	2	12	27
	200ppm	1	1	2	4	43	0	8	0	47
		2	17	1	23	12	1	4	6	31
		3	11	2	3	0	1	3	7	10
	400ppm	1	0	3	8	14	0	3	1	23
		2	17	4	13	1	2	3	4	17
		3	0	1	1	0	6	0	6	6

¹Time intervals 1-3 represent, respectively, 0-40, 41-60, 61-80 weeks.

²With tumor; ³Without tumor; ⁴Insufficient data

Table 2: Likelihood contribution of each event

Event	Likelihood contribution	Count for interval j
Death from fatal tumor	$Q(t_{j-1})\{P(t_{j-1}) - P(t_j)\}$	d_j
Death with incidental tumor	$\{Q(t_{j-1}) - Q(t_j)\}\{P(t_j) - S(t_j)\}$	a_{1j}
Death without tumor	$S(t_j)\{Q(t_{j-1}) - Q(t_j)\}$	b_{1j}
Sacrifice with tumor	$Q(t_j)\{P(t_j) - S(t_j)\}$	a_{2j}
Sacrifice without tumor	$S(t_j)Q(t_j)$	b_{2j}

Table 3: Counts for fitting the discrete model to data and estimated quantities from the Benzidine Data

Group	j^1	ad_j^2	b_{1j}	a_{2j}	b_{2j}	$\hat{\pi}_j$	\hat{p}_j	\hat{q}_j	$\hat{S}(t_j)$	$\hat{P}(t_j)$	$\hat{Q}(t_j)$	Tb_j	To_j	\hat{d}_j^3	\tilde{d}_j^4
F1F	1	0	3	21	170	0.89	1.00	0.99	0.89	1.00	0.99	54	54	0	0
	2	39	2	64	84	0.57	0.88	0.99	0.50	0.88	0.98	147	114	38	33
	3	45	8	47	19	0.29	0.79	0.70	0.20	0.69	0.69	92	47	25	40
F1M	1	1	4	3	116	0.98	1.00	0.99	0.97	1.00	0.99	9	9	0.9	1
	2	4	9	23	85	0.79	0.99	0.95	0.78	0.99	0.93	45	39	2	4
	3	10	3	30	46	0.61	0.91	0.94	0.55	0.90	0.88	40	23	8	9
F2F	1	1	14	14	171	0.93	1.00	0.97	0.93	1.00	0.97	37	37	0	0
	2	40	10	61	77	0.56	0.89	0.93	0.50	0.89	0.91	141	119	32	29
	3	43	9	39	21	0.35	0.77	0.70	0.24	0.69	0.63	82	42	26	42
F2M	1	0	22	1	184	1.00	1.00	0.96	1.00	1.00	0.96	2	2	0	0
	2	4	11	21	128	0.86	0.99	0.96	0.85	0.99	0.91	41	40	2	3
	3	13	8	29	77	0.73	0.92	0.91	0.66	0.91	0.83	42	26	10	11

¹Time intervals 1-3 represent, respectively, 0-40, 41-60, 61-80 weeks.

² $a_{1j}+d_j$

³Estimated number of fatal tumors

⁴Number of fatal tumors assigned by pathologists

Table 4: Estimated time-to-tumor-onset survival functions $\hat{S}(t)$ from (1) and $\tilde{S}(t)$ by the method of Kodell and Ahn (1997), for the benzidine data.

Interval j^a	F1 Female		F1 Male		F2 Female		F2 Male	
	$\hat{S}(t_j)$	$\tilde{S}(t_j)$	$\hat{S}(t_j)$	$\tilde{S}(t_j)$	$\hat{S}(t_j)$	$\tilde{S}(t_j)$	$\hat{S}(t_j)$	$\tilde{S}(t_j)$
1	0.892	0.891	0.972	0.972	0.925	0.925	0.995	0.995
2	0.499	0.499	0.779	0.779	0.498	0.498	0.853	0.853
3	0.199	0.199	0.545	0.545	0.239	0.239	0.664	0.664

^aTime intervals 1-3 represent, respectively, 0-40, 41-60, 61-80 weeks

Table 5: Estimated lag time (in weeks) in (7) for the Benzidine data

Strain	Sex	Concentration (ppm)	RL ¹	Proposed
F1	Female	60	17.5	20.9
		120	20.5	
		200	20.2	
		400	21.1	
	Male	60	* ²	16.9
		120	28.3	
		200	24.8	
		400	30.4	
F2	Female	60	20.3	21.4
		120	21.7	
		200	18.0	
		400	21.8	
	Male	60	14.9	26.3
		120	20.7	
		200	30.9	
		400	19.3	

¹Estimated lag time using the method by Kodell and Nelson (1980)

²Insufficient data to obtain estimates

Table 6: The values of $t\text{lag}_r$ in (6) and $\text{lag}_r(r)$ used for the calculation of $t\text{lag}_r$ in each group for the benzidine data

Group	F1 female	F1 male	F2 female	F2 male
$t\text{lag}_1$	34.897	23.515	35.713	42.747
$t\text{lag}_2$	8.673	15.386	9.285	14.561
$t\text{lag}_3$	5.000	5.000	5.000	5.000
$\text{lag}_1(1)$	0.000	0.897	0.000	0.000
$\text{lag}_2(2)$	8.429	0.183	4.965	0.072
$\text{lag}_3(3)$	2.733	0.413	1.985	0.126

Table 7: Average of the estimated and actual lag times for simulated data with sacrifices at the end of each NTP intervals. Standard error estimates are given in parentheses. For each configuration, the average is taken from 5000 trials. The competing risks survival rate is 0.5.

Background Tumor rate	Tumor lethality	Average lag time	
		Estimated	Actual
0.05	≤ 0.05	21.59 (0.21)	25.85
	$\cong 0.35$	16.54 (0.11)	17.63
	$\cong 0.60$	13.45 (0.08)	11.82
0.15	≤ 0.05	26.87 (0.22)	27.60
	$\cong 0.35$	18.49 (0.12)	18.29
	$\cong 0.60$	15.14 (0.09)	12.44
0.30	≤ 0.05	29.87 (0.20)	28.23
	$\cong 0.35$	20.13 (0.10)	19.11
	$\cong 0.60$	16.65 (0.08)	13.17

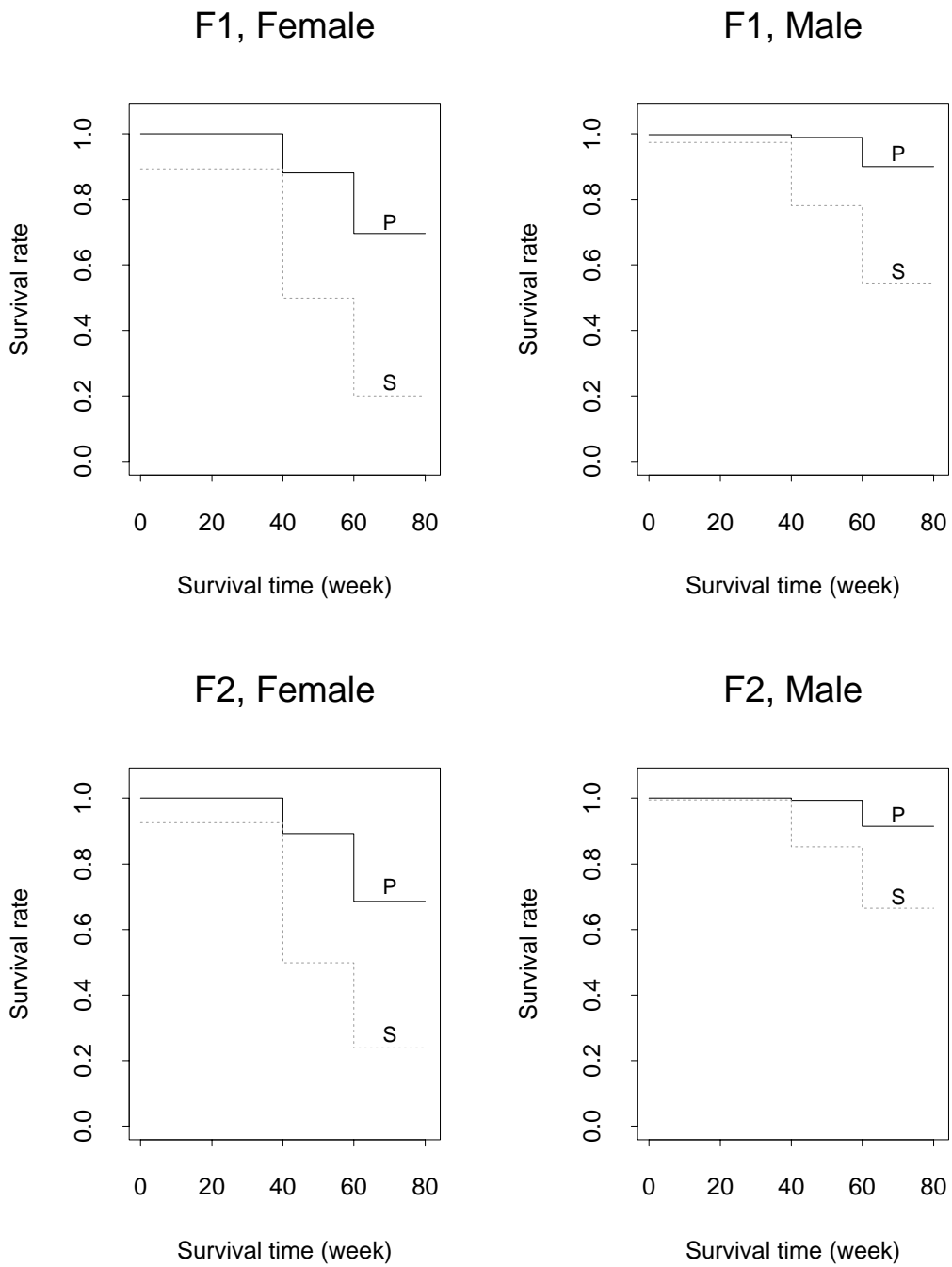


Figure 1. Time-to-tumor-onset survival functions (S) and tumor-specific survival functions (P) using the proposed method for the Benzidine data.