

A dose-response test via closed-form solutions for constrained  
MLEs in survival/sacrifice experiments

Wonkuk Kim and Hongshik Ahn\*

Department of Applied Mathematics and Statistics  
Stony Brook University, Stony Brook, NY 11794-3600, U.S.A.

\*Phone: (631)632-8372, Fax: (631)632-8490, e-mail: hahn@ams.stonybrook.edu

Hojin Moon

Division of Biometry and Risk Assessment  
National Center for Toxicological Research, FDA  
Jefferson, AR 72079, U.S.A.

Short title: A dose-response test via closed-form solutions

Corresponding author: Hongshik Ahn

## SUMMARY

In most survival-sacrifice experiments in animal carcinogenicity studies, the onset of the tumor of interest is not clinically observable. Due to the complexity of constraints for a biological justification, recently developed methods for estimating the tumor onset function and tumor-specific survival function employ computer-intensive numerical solutions. In this paper, closed-form solutions for nonparametric maximum likelihood estimators are derived under explicit and implicit inequality constraints obtained from the monotonicity of the survival functions. Our methods do not require cause-of-death information. The proposed methods can be used to estimate the tumor onset function and the survival function of the tumor of interest. We use the proposed estimators for the development of our new dose-response trend test. A modification of the Poly- $k$  test is made by replacing the time-at-risk weight to a function of the tumor onset survival function. The weighted least square regression method is applied to the estimated survival functions in order to construct a dose-response trend test. A simulation study is conducted to evaluate the performance of the proposed test and compare it with existing trend tests. A real example is used to illustrate the methods.

Key Words: Animal carcinogenicity experiments; Competing risks; Dose response; Survival function; Tumor onset.

## 1 INTRODUCTION

The main goal of long-term animal carcinogenicity experiments is to test the statistical significance of a dose-response relationship among treatment groups. When the tumor onset time is observable, standard lifetable methods can be applied to test the differences in tumor incidence rates [1]. However, it is more complicated when information concerning onset of specific tumors is confounded with information concerning the effect of the other causes on mortality.

Various methods have been proposed for testing dose-related trend. Since the estimation of incidence rates for a certain tumor can be affected by competing causes of death, a number of analyses require cause-of-death (COD) information [2, 3, 4, 5, 6]. However, pathologists often claim that accurate determinations of the cause of death are impossible, and classification errors can cause biases [7, 8]. Peto [2] proposed an approach that requires COD information for each animal to be determined by pathologists. Bailer and Portier [9] proposed the Poly-3 test which does not require the COD information. The Poly-3 test is derived by making an adjustment to

the Cochran-Armitage (CA) test [10, 11] by using a fractional weighting scheme for animals not at full risk of tumor development. Bieler and Williams [12] modified the Poly-3 test by adjusting the variance of the test statistic.

The Poly-3 test improves the sensitivity of the CA test caused by the presence of treatment-induced mortality unrelated to the tumor of interest. Performance of the Poly-3 test depends on how closely it represents the correct specification of the time-at-risk weight in the data. The STP Peto Analysis Working Group [13, 14, 15] pointed out the limitation of the Poly- $k$  test due to the lack of reliable knowledge of the appropriate value of  $k$  to use, and the limitation of the Peto test due to the uncertainty in pathologist-assigned COD.

Moon et al. [16] developed a method for estimating  $k$  in the Poly- $k$  test by equating the empirical estimator of the lifetime cumulative tumor rate based on the Poly- $k$  fractional weighting scheme to a separately estimated lifetime cumulative tumor rate based on the method by Kodell and Ahn [17]. This is done by developing methods for estimating the tumor onset distribution by deriving the estimated lifetime cumulative tumor incidence rates. Moon et al. [18] investigated an empirical distribution of the Poly-3 test statistic using an age-adjusted bootstrap-based method and compared it with the Poly-3 test statistic referenced to the assumed standard normal distribution.

In this paper, we propose a dose-response trend test based on the time-at-risk weight, but without using the shape parameter  $k$ . A modification of the Poly- $k$  test is made by replacing the time-at-risk weight to a function of the tumor onset survival function. The proposed method uses maximum likelihood estimators (MLEs) derived in this study for the tumor onset survival function. We apply the weighted least square regression method to the estimated survival functions in order to estimate the time-at-risk weight without assuming a Weibull distribution on tumor onset time.

Due to the complexity of constraints, recently developed methods for estimating the tumor onset function and tumor-specific survival function employ computer-intensive numerical solutions [19, 20]. Several papers address estimation of the tumor incidence rates or survival functions. Williams and Portier [21] proposed analytic expressions for nonparametric MLEs of the tumor incidence rate for designs with one or two interim sacrifices in addition to the terminal sacrifice. Williams and Portier [22] later imposed a boundary condition for a nonnegative tumor incidence rate for up to three sacrifices. As mentioned in their paper, the estimator of the tumor incidence rate in [21] is analytically equivalent to that proposed by Malani and Van Ryzin [23]. Ahn, Moon

and Kodell [19] developed a different constrained nonparametric maximum likelihood estimation method for attribution of tumor lethality in the absence of COD information by estimating the tumor onset function and tumor-specific survival functions.

We develop closed-form solutions for nonparametric MLEs from the multinomial likelihood function in this paper based on the mixture of fatal and incidental tumors under inequality constraints given in [19, 20]. We derive the large-sample variance of the estimated tumor onset survival function. This analytical method can replace the existing complicated computer-intensive method, and provide unique global MLEs which may not be guaranteed by a numerical method. These closed-form solutions are used to modify the Poly- $k$  test to a trend test without using the parameter  $k$ . The proposed method estimates the tumor onset function, tumor-specific survival function and survival function for competing risks for each group. Using these estimators of the survival functions and the variance estimators, the time-at-risk weight for the proposed dose-response trend test can be obtained.

## 2 ANALYTICAL SOLUTIONS FOR CONSTRAINED MLEs

For maximum likelihood estimation, we assume a single group of animals in an animal carcinogenicity experiment. Let  $0 = t_0 < t_1 < \dots < t_s$  be  $s + 1$  appropriately spaced time points, where  $t_s$  denotes the terminal sacrifice time. Suppose that  $N_0$  animals are followed over time for the development of irreversible and occult tumors. We assume that all animals come from the same population and are born without tumor at the beginning of the experiment.

Let  $X$ ,  $Y$  and  $Z$  be nonnegative random variables representing the tumor onset time, death time from the tumor, and death time from competing risks in animal carcinogenicity studies with multiple-sacrifices, respectively. We assume that  $Y$  is dependent on  $X$  such that  $Y \geq X$ . Let  $Z$  be independent to  $X$  and  $Y$ . Consider three survival functions:  $S(t) = \Pr(X > t)$ ,  $F(t) = \Pr(Y > t)$ , and  $G(t) = \Pr(Z > t)$ . Consider the  $i$ th of the  $N_0$  independent individuals following  $(X_i, Y_i, Z_i)$ . Suppose randomly chosen animals are pre-assigned to be sacrificed and leave the experiment at some  $t_j$  in a finite set  $\{t_0 = 0, \dots, t_s = t_{\max}\}$ . If the  $i$ th animal is supposed to be censored at  $t_j$ , then we define a random variable  $C_i = t_j$ .

To develop the likelihood contributions, we consider five key events and the number of animals

for each event. For the  $j$ th interval in a given experimental group, let  $d_j$  denote the number of animals observed to die from fatal tumors. Let  $a_{1j}$  denote the number of animals observed to die with incidental tumors and  $b_{1j}$  the number observed to die without tumors. Let  $a_{2j}$  and  $b_{2j}$  denote the number of sacrificed animals with and without tumor of interest, respectively.

Assuming that the cause of death is not available, four types of events are observable of which counts can be expressed as  $ad_j = \#\{i|t_{j-1} < \min(Y_i, Z_i) \leq t_j \leq C_i, X_i \leq \min(Y_i, Z_i)\}$  by combining the first two events,  $b_{1j} = \#\{i|t_{j-1} < Z_i \leq t_j \leq C_i, Z_i \leq X_i\}$ ,  $a_{2j} = \#\{i|X_i \leq C_i = t_j < \min(Y_i, Z_i)\}$ , and  $b_{2j} = \#\{i|C_i = t_j < X_i, C_i < Z_i\}$ .

The likelihood function can be formulated using the likelihood contributions of Moon et al. [20] as follows:

$$L \propto \prod_{j=1}^s [G(t_j)\{F(t_j) - S(t_j)\}]^{a_{2j}} [S(t_j)G(t_j)]^{b_{2j}} [G(t_{j-1})\{F(t_{j-1}) - F(t_j)\} + \{G(t_{j-1}) - G(t_j)\}\{F(t_j) - S(t_j)\}]^{ad_j} [S(t_j)\{G(t_{j-1}) - G(t_j)\}]^{b_{1j}}.$$

After reparameterization of  $f_j = F(t_j)/F(t_{j-1})$ ,  $g_j = G(t_j)/G(t_{j-1})$ , and  $\pi_j = S(t_j)/F(t_j)$ , the log-likelihood function is derived as

$$l = \sum_{j=1}^s \left[ a_{2j} \{\log f_j + \log g_j + \log(1 - \pi_j)\} + b_{2j} (\log f_j + \log g_j + \log \pi_j) + N_j (\log f_j + \log g_j) + ad_j \log\{1 - f_j + (1 - g_j)f_j(1 - \pi_j)\} + b_{1j} \{\log f_j + \log(1 - g_j) + \log \pi_j\} \right] + \text{constant},$$

where  $N_j$  denotes the number of individuals at  $t_j$ .

For the  $j$ th interval, there are three monotonicity conditions on this problem:

$$0 \leq f_j \leq 1, 0 \leq g_j \leq 1 \text{ and } \pi_j f_j \leq \pi_{j-1} \leq 1. \quad (1)$$

Define  $N_j^*$  as the number of individuals in the experiment at  $t_{j-}$ , which is right before censoring at the end of the  $j$ th interval. Then  $N_j^* = N_j + a_{2j} + b_{2j}$  and  $N_{j-1} = N_j^* + ad_j + b_{1j}$ . The critical

points of the log-likelihood function can be obtained by solving the following three equations:

$$\frac{\partial l}{\partial f_j} = -\frac{ad_j\{g_j + (1 - g_j)\pi_j\}}{1 - f_j\{g_j + (1 - g_j)\pi_j\}} + \frac{N_j^* + b_{1j}}{f_j} = 0 \quad (2)$$

$$\frac{\partial l}{\partial g_j} = -\frac{ad_j f_j(1 - \pi_j)}{1 - f_j\{g_j + (1 - g_j)\pi_j\}} + \frac{N_j^*}{g_j} - \frac{b_{1j}}{1 - g_j} = 0 \quad (3)$$

$$\frac{\partial l}{\partial \pi_j} = -\frac{ad_j f_j(1 - g_j)}{1 - f_j\{g_j + (1 - g_j)\pi_j\}} + \frac{b_{1j} + b_{2j}}{\pi_j} - \frac{a_{2j}}{1 - \pi_j} = 0 \quad (4)$$

for  $j = 1, \dots, s$ . For convenience, set  $\pi_{s+1}f_{s+1} = 0$ . The constrained MLEs are obtained from the following theorem.

**Theorem 2.1** *If  $\pi_j \in [\pi_{j+1}f_{j+1}, 1]$ ,  $f_j \in [0, 1]$ , and  $g_j \in [0, 1]$ ,  $j = 1, \dots, s$ , then the MLEs of  $f_j, g_j$  and  $\pi_j$  are obtained as follows.*

*If  $b_{2j}/(a_{2j} + b_{2j}) \leq b_{1j}/(ad_j + b_{1j})$ , then the MLEs are*

$$\begin{aligned} \hat{\pi}_j &= \max \left\{ \frac{b_{1j} + b_{2j}}{N_{j-1} - N_j}, \hat{\pi}_{j+1} \hat{f}_{j+1} \right\} \\ \hat{g}_j &= \frac{N_j^*}{N_{j-1}} \\ \hat{f}_j &= 1. \end{aligned}$$

*If  $b_{2j}/(a_{2j} + b_{2j}) \geq b_{1j}/(ad_j + b_{1j})$ , then the MLEs are*

$$\begin{aligned} \hat{\pi}_j &= \max \left\{ \frac{b_{2j}}{a_{2j} + b_{2j}}, \hat{\pi}_{j+1} \hat{f}_{j+1} \right\} \\ \hat{g}_j &= \frac{N_j^* \hat{\pi}_j}{N_j^* \hat{\pi}_j + b_{1j}} \\ \hat{f}_j &= \frac{N_j^* + b_{1j}}{N_{j-1} \{\hat{g}_j + (1 - \hat{g}_j) \hat{\pi}_j\}}. \end{aligned}$$

A proof of the above theorem is given in the Appendix.

We now consider the large sample variance estimation of  $S(t_j)$ . We estimate  $S(t)$  at  $t = t_j$  by products  $S(t_j) = (f_1 \pi_1 / \pi_0) \cdots (f_{j-1} \pi_{j-1} / \pi_{j-2}) (f_j \pi_j / \pi_{j-1})$  of the  $f_i$ 's and  $\pi_i$ 's. The variance

estimation of the survival function of  $X$  from Greenwood's formula [24] may be written as

$$\widehat{\text{var}} \left[ \hat{S}(t_j) \right] = \hat{S}^2(t_j) \left[ \sum_{i=1}^j \frac{1}{\hat{f}_i^2} \widehat{\text{var}}(\hat{f}_i) + \frac{1}{\hat{\pi}_j^2} \widehat{\text{var}}(\hat{\pi}_j) + \frac{2}{\hat{f}_j \hat{\pi}_j} \widehat{\text{cov}}(\hat{f}_j, \hat{\pi}_j) \right],$$

where  $\widehat{\text{var}}(\hat{f}_i)$ ,  $\widehat{\text{var}}(\hat{\pi}_j)$  and  $\widehat{\text{cov}}(\hat{f}_j, \hat{\pi}_j)$  can be obtained from the observed information matrix. Note that  $\hat{S}(t_j)$ ,  $\hat{f}_i$  and  $\hat{\pi}_i$  are rational functions of observed counts in the above Theorem. It is easy to see from simple algebra that this variance estimation is equivalent to a binomial variance estimation if the set of prescribed censoring time points consists of a single time point  $t_1$  and  $b_{11} = 0$ . That is,

$$\widehat{\text{var}} \left[ \hat{S}(t_1) \right] = \hat{S}^2(t_1) \left[ \frac{1 - \hat{f}_1}{N_0 \hat{f}_1} + \frac{1 - \hat{\pi}_1}{(a_{21} + b_{21}) \hat{\pi}_1} \right] = \frac{\hat{S}(t_1) \{1 - \hat{S}(t_1)\}}{N_0}.$$

### 3 DOSE-RESPONSE TREND TEST

The Cochran-Armitage (CA) trend test [10, 11] can be viewed as a weighted least square regression across dose groups in which a binomial variance is assumed for the quantal response estimates of lifetime tumor incidence. Suppose that the  $i$ th dose group initially consists of  $n_i$  animals. Let  $y_{ij}$  be an indicator function of tumor status of the  $j$ th animal in the  $i$ th group. The test statistic is given as

$$Z_{CA} = \frac{\sum n_i (p_i - \bar{p}_w) (d_i - \bar{d}_w)}{\sqrt{pq \sum n_i (d_i - \bar{d}_w)^2}},$$

where  $p_i = y_i/n_i = \sum_j y_{ij}/n_i$ ,  $\bar{p}_w = \sum w_i p_i / \sum w_i$ ,  $\bar{d}_w = \sum w_i d_i / \sum w_i$ ,  $w_i = n_i/(pq)$  and  $p = \sum n_i p_i / n_i$ . The test assumes that all animals are at equal risk of developing a tumor during the study under the null hypothesis of equal tumor incidence rate across dose groups. However, some animals die without tumor of interest and leave the experiment before the study is terminated. It may be considered as right censoring.

Bailer and Portier [9] introduced a modification to the CA test by accounting for mortality of animals in adjusting the denominator in the risk set. Through a weighting scheme, the test modifies the denominator  $n_i$  of the proportion  $p_i$  of tumor incidence in each group by allowing fractional counts for animals not at full risk. The Poly- $k$  test statistic is obtained by replacing  $n_i$  with  $n'_i = \sum \alpha_{ij}$ . Here,  $\alpha_{ij} = 1$  if the  $j$ th animal in the  $i$ th group is at full risk; died with tumor

or survived to the end of the experiment. A weight  $\alpha_{ij} = (t_{ij}/t_{\max})^k$  is given if the animal died or was sacrificed without tumor at  $t = t_{ij}$ .

Later, Bieler and Williams [12] derived a ratio statistic from estimating the variance of  $p'_i = y_i/n'_i$  under  $H_0$  with the first order Taylor series approximation which may be written as

$$\text{var}_0(p'_i) = \frac{n_i}{(n'_i)^2} \frac{\sum_{i,j} (r_{ij} - \bar{r}_i)^2}{n - (I + 1)},$$

where  $I + 1$  is the total number of groups in the experiment including a control group, and  $r_{ij} = y_{ij} - p'_i \alpha_{ij}$ . The modified test statistic is

$$Z_r = \frac{\sum a_i (p'_i - \bar{p}'_w) (d_i - \bar{d}_w)}{\sqrt{C \sum a_i (d_i - \bar{d}_w)^2}}, \quad (5)$$

where  $a_i = (n'_i)^2/n_i$  and  $C = [n - (I + 1)]^{-1} \sum_{i,j} (r_{ij} - \bar{r}_i)^2$  so that  $\text{var}_0(p'_i) = C/a_i$ .

In this paper, we propose a dose-response trend test statistic that does not include the shape parameter  $k$ , but use the maximum likelihood solutions for  $S(t_j)$ . In the estimation of  $S(t)$  used for testing, the control and dose groups are pooled for estimating the parameters, since the estimates are more stable with the pooled sample. It is common to pool the data in order to obtain a stable estimation [12]. If the sample sizes are big enough in the control group, we may use this group alone to estimate the baseline survival function  $S(t_j)$ . A simple modification to the Poly- $k$  test may be obtained as follows:

$$\alpha_{ij} = \frac{[1 - \hat{S}(t_m)]}{\hat{S}(t_m)} \bigg/ \frac{[1 - \hat{S}(t_{\max})]}{\hat{S}(t_{\max})} \quad (6)$$

for an animal which died without tumor at  $t_{ij} \in ((t_{m-1} + t_m)/2, (t_m + t_{m+1})/2]$ . This includes death by sacrifice. Otherwise, we set  $\alpha_{ij} = 1$ . The approach of breaking the time period into these intervals rather than the intervals defined in Section 2 is used for calculating the above weight rather than the interval defined in Section 2 in order to better reflect the survival time of an animal. Note that Equation (6) for the weights was derived by Bailer and Portier [9], but they assumed a Weibull distribution for time to tumor onset because the estimation of  $S(t)$  was not available. We are now able to estimate the survival functions at the prescribed sacrifice time using the constrained MLEs

of the parameters obtained in Section 2. We use the test statistic (5) with the estimated weights shown in (6).

## 4 SIMULATION STUDY

A typical bioassay design with four dose groups of 50 animals each and an experimental duration of two years is used in this study according to standard designs of the National Toxicology Program (NTP), with interim sacrifice times at 52, 78, 92 weeks, and a terminal sacrifice at 104 weeks [25, 26]. In each group, six animals are randomly selected to be sacrificed at the end of each interval. All the remaining live animals are sacrificed at the end of the experiment. The dose levels used in the simulation are 0, 1, 2 and 4. These dose levels are used based on typical carcinogenicity studies. In real data analysis, actual dose levels are supposed to be used.

Distributions of time to onset and time to death were of the form used by Portier et al. [25]. It is assumed that three independent random variables  $T_1$  (time to tumor onset),  $T_2$  (time from onset until death from the tumor), and  $T_3$  (time until death from a competing risk) completely determine the observed outcome for each animal. The survival function of  $T_1$  for the  $i$ th group is modeled as the one used in [16] such that

$$S_i(t) = \exp \left[ -\delta(\ell_i) \left( \frac{t}{t_{\max}} \right)^k \right], \quad (7)$$

where  $t_{\max}$  represents the duration of the study or the time for the terminal sacrifice and  $\ell_i$  is the dose level for the  $i$ th group. The value of  $k$  is set to be 1.5, 3 or 6 for the Weibull tumor onset distribution with the shape parameter of 1.5, 3 or 6, respectively. The best performance is expected from the Poly- $k$  test when the correct  $k$  is used for the data with the tumor onset survival function (7). The value of  $\delta(\ell_i)$  is chosen such that the probability of tumor onset by the end of the experiment attains the desired rate. The tumor rates are chosen to be either 0.05, 0.15 or 0.30 for the control group. The tumor rates are set to be the same across dose groups for size evaluations. For power comparisons, the tumor rates at the highest dose group by 104 weeks are set to be 5, 3, and 2 times higher than the background tumor rates of 0.05, 0.15 and 0.30, respectively. The low and intermediate doses are chosen to be a quarter and a half of the highest dose, respectively. This simulation design is for frequent tumors. For spontaneous tumors with low rates, exact tests can be applied [27].

The survival function for  $T_3$  is modeled as  $Q(t) = \exp[-\phi(\gamma_1 t + \gamma_2 t^{\gamma_3})]$ , where  $\phi \geq 1$ , and  $\gamma_1, \gamma_2$ , and  $\gamma_3$  are nonnegative. The same parameter values as in Moon et al. [16] are used in our simulation. The competing risks survival rate (CRSR) considered in this simulation is (0.7, 0.6, 0.5, 0.4), (0.5, 0.4, 0.3, 0.2), (0.6, 0.6, 0.6, 0.2) or (0.6, 0.6, 0.6, 0.6) for the control and three dose groups. These combinations of the CRSR are considered in order to represent most of the actual animal tumor experiments seen in practice [18].

The survival distribution for tumor-induced mortality,  $T_2$ , has the same form as the one for death from competing risks, and the values of  $\gamma_1, \gamma_2$  and  $\gamma_3$  remain the same. The parameter  $\phi$  is chosen to reflect various tumor lethality. For each configuration in our simulation study, 10,000 simulated data sets were generated and a nominal significance level of  $\alpha = 0.05$  was used.

Tables 1 and 2 display the probability of a Type I error and power for the configurations considered in this simulation study. When CRSR is the same across dose groups, the test results do not appear to be affected by the choice of  $k$  because the inter-current mortality patterns are equal across dose groups as addressed in Moon et al. [18]. Thus we focus on the configurations with different CRSR across dose groups.

The performance of our method is comparable to that of the Poly- $k$  test with the fixed correct value of  $k$ . Our method outperforms the Poly- $k$  test using a fixed, but incorrect, value of  $k$  in the sense of controlling the Type I error rate. Although the Poly-3 test yields reasonable size for data with  $k = 3$ , it tends to be anticonservative for data with  $k = 1.5$  and too conservative for data with  $k = 6$ . Even for data with  $k = 3$ , the proposed method gave comparable results to the Poly-3 test which is suitable for this tumor onset distribution.

For all the configurations considered in these simulations, our method appears to be quite robust. The Type I error rate obtained from our method ranges from 2.37 to 6.04, while those for the CA test and the Poly- $k$  tests with fixed  $k$  at 1.5, 3, and 6 range within (0.63, 5.90), (1.26, 5.92), (1.97, 6.51) and (3.89, 10.95), respectively.

For data from the Weibull tumor onset distribution with  $k = 6$  and a tumor rate of 30%, the widely used Poly-3 test is quite conservative and could result in a significant loss of power (see Table 2). For data with  $k = 1.5$  and a tumor rate of 0.05, the Poly-3 test shows more inflated size compared to our method. The proposed test closely follows the Poly- $k$  test with a true value of  $k = 1.5$ . The size and power of the proposed method show a substantial improvement over the

Table 1: Simulated size (%) corresponding to the nominal 5% significance level for the proposed test, CA test, the Poly- $k$  tests with several fixed  $k$  values, and the Poly- $k$  with estimated  $k$  [16]. All tests are based on a dose scaling of 0, 1, 2, 4 with 50 animals per group and 10,000 trials for each configuration; approximate lethality is 35%.

$k_{true}$	CRSR	Tumor Rate	Poly- $k$ with fixed $k$			Poly- $k$ with $\hat{k}$	Proposed test	
			CA	$k = 1.5$	$k = 3$			$k = 6$
1.5	(.7 .6 .5 .4)	.05	5.38	5.89	6.37	7.34	5.77	6.04 (0.24)
		.15	3.86	5.14	6.09	7.82	5.57	5.04 (0.22)
		.3	3.46	4.83	6.02	8.41	4.88	4.53 (0.21)
	(.5 .4 .3 .2)	.05	4.91	5.46	6.30	7.69	5.76	5.81 (0.23)
		.15	3.34	4.79	6.21	9.03	5.24	4.86 (0.22)
		.3	3.46	4.88	6.51	10.17	4.39	4.47 (0.21)
	(.6 .6 .6 .2)	.05	4.35	4.98	5.87	7.34	5.65	5.42 (0.23)
		.15	3.06	4.64	6.35	9.30	5.08	4.74 (0.21)
		.3	2.90	4.82	7.03	10.95	4.25	4.46 (0.21)
	(.6 .6 .6 .6)	.05	5.76	5.78	5.85	5.84	5.70	5.72 (0.23)
		.15	5.04	5.45	5.42	5.17	5.60	5.45 (0.23)
		.3	5.22	5.38	5.24	5.08	4.70	5.33 (0.22)
3.0	(.7 .6 .5 .4)	.05	4.19	4.49	5.02	5.87	5.36	4.99 (0.22)
		.15	3.13	4.35	5.02	6.51	5.20	4.58 (0.21)
		.3	2.33	3.36	4.33	6.51	4.63	3.73 (0.19)
	(.5 .4 .3 .2)	.05	4.22	4.90	5.69	6.73	5.25	5.69 (0.23)
		.15	2.60	4.07	5.15	7.40	4.48	4.79 (0.21)
		.3	2.11	3.41	4.80	7.75	3.90	4.09 (0.20)
	(.6 .6 .6 .2)	.05	3.70	4.27	5.06	6.21	4.72	5.00 (0.22)
		.15	2.25	3.46	4.66	6.97	4.03	4.13 (0.20)
		.3	1.54	2.84	4.42	7.53	3.48	3.28 (0.18)
	(.6 .6 .6 .6)	.05	5.35	5.34	5.44	5.45	5.52	5.43 (0.23)
		.15	4.98	5.27	5.39	5.43	5.45	5.36 (0.23)
		.3	5.42	5.62	5.59	5.48	4.93	5.66 (0.23)
6.0	(.7 .6 .5 .4)	.05	4.07	4.47	4.85	5.28	4.84	5.06 (0.22)
		.15	2.35	3.29	3.90	5.07	4.11	4.11 (0.20)
		.3	1.51	2.21	2.89	4.25	4.02	3.20 (0.18)
	(.5 .4 .3 .2)	.05	3.30	3.65	4.01	4.73	4.70	4.48 (0.21)
		.15	1.91	2.71	3.24	4.85	3.79	3.80 (0.19)
		.3	1.05	1.67	2.47	4.05	3.07	2.93 (0.17)
	(.6 .6 .6 .2)	.05	2.59	3.10	3.54	4.12	3.90	3.88 (0.19)
		.15	1.23	1.88	2.64	4.18	3.13	3.17 (0.18)
		.3	0.63	1.26	1.97	3.87	2.73	2.37 (0.15)
	(.6 .6 .6 .6)	.05	5.90	5.92	5.97	5.91	5.66	5.95 (0.24)
		.15	4.72	5.16	5.38	5.46	5.45	5.28 (0.22)
		.3	4.45	4.83	4.93	4.91	4.97	5.04 (0.22)

Table 2: Simulated power (%) corresponding to the nominal 5% significance level for the proposed test, CA test, the Poly- $k$  tests with several fixed  $k$  values, and the Poly- $k$  with estimated  $k$  [16]. All tests are based on a dose scaling of 0, 1, 2, 4 with 50 animals per group and 10,000 trials for each configuration; approximate lethality is 35%.

$k_{true}$	CRSR	Tumor Rate	Poly- $k$ with fixed $k$			Poly- $k$ with $\hat{k}$	Proposed test	
			CA	$k = 1.5$	$k = 3$			$k = 6$
1.5	(.7 .6 .5 .4)	.05	86.6	89.2	90.2	91.8	88.5	88.8 (0.32)
		.15	92.1	93.9	95.0	96.3	93.6	93.6 (0.24)
		.3	85.8	90.0	91.8	93.9	89.3	89.0 (0.31)
	(.5 .4 .3 .2)	.05	84.2	87.6	89.4	91.7	86.3	87.4 (0.33)
		.15	90.5	93.3	94.9	96.5	91.8	92.7 (0.26)
		.3	83.6	88.6	91.1	93.7	86.7	87.4 (0.33)
	(.6 .6 .6 .2)	.05	82.3	86.7	88.9	91.4	85.9	86.3 (0.34)
		.15	89.7	93.3	94.8	96.7	91.5	92.4 (0.26)
		.3	81.2	88.0	91.3	94.4	86.5	86.4 (0.34)
	(.6 .6 .6 .6)	.05	88.5	89.1	89.2	89.0	89.1	89.1 (0.31)
		.15	94.6	94.9	94.7	94.4	94.4	94.8 (0.22)
		.3	91.0	91.5	91.1	90.0	90.6	91.5 (0.28)
3.0	(.7 .6 .5 .4)	.05	78.9	82.9	84.4	86.7	82.6	83.3 (0.37)
		.15	84.9	88.3	90.4	92.5	89.1	89.0 (0.31)
		.3	73.7	80.8	84.4	88.3	83.1	81.9 (0.39)
	(.5 .4 .3 .2)	.05	73.9	78.5	80.9	84.7	77.4	79.8 (0.40)
		.15	79.8	85.2	88.5	92.3	84.4	86.2 (0.34)
		.3	68.3	77.1	82.2	87.6	76.5	78.7 (0.41)
	(.6 .6 .6 .2)	.05	70.6	76.7	79.9	84.0	76.7	78.6 (0.41)
		.15	76.8	83.5	87.6	91.5	83.9	84.7 (0.36)
		.3	63.0	74.8	81.1	87.8	76.3	76.6 (0.42)
	(.6 .6 .6 .6)	.05	84.9	85.6	85.8	85.8	85.5	85.6 (0.35)
		.15	91.6	92.2	92.3	92.3	91.5	92.2 (0.27)
		.3	85.5	87.0	87.3	86.9	86.9	87.3 (0.33)
6.0	(.7 .6 .5 .4)	.05	67.6	71.9	73.6	76.8	74.4	74.6 (0.44)
		.15	70.6	76.4	79.9	83.9	81.1	80.3 (0.40)
		.3	55.2	64.4	70.2	77.6	73.0	71.5 (0.45)
	(.5 .4 .3 .2)	.05	57.3	61.7	64.9	70.1	64.6	66.5 (0.47)
		.15	59.0	66.3	71.3	78.3	71.2	72.9 (0.44)
		.3	42.8	52.7	60.2	70.9	61.1	62.8 (0.48)
	(.6 .6 .6 .2)	.05	52.2	58.7	62.9	68.3	62.7	64.2 (0.48)
		.15	52.2	62.3	68.9	76.9	69.8	70.5 (0.46)
		.3	34.5	47.5	57.4	69.2	60.4	60.0 (0.49)
	(.6 .6 .6 .6)	.05	78.4	79.4	79.9	80.2	79.8	80.1 (0.40)
		.15	85.5	86.6	87.2	87.4	86.4	87.1 (0.34)
		.3	77.7	80.5	81.6	82.2	81.2	81.7 (0.39)

Poly-3 test.

The proposed test has slightly higher power than Poly-1.5 and Poly-3 tests when the actual value of  $k$  is 6. When the actual  $k$  is 1.5, 3 or 6, the power of the proposed test is about the same or slightly lower than the Poly-1.5, Poly-3 and Poly-6 tests. This result shows the same tendency as the size of the proposed test which shows reasonable control of the Type I error rate.

## 5 EXAMPLE

In rodents a calorie-restricted diet compared to ad libitum feeding markedly decreases tumor incidence and increases lifespan [28, 29, 30, 31]. Data from the Project on Calorie Restriction (PCR) conducted at the National Center for Toxicological Research are analyzed to investigate the effects of calorie restriction on tumorigenicity and longevity in Fischer 344 rats [32]. The data contain the lifetimes, tumors found, gender and treatment group information such as *ad libitum* (AL) or calorie restriction (CR). The study involved up to six scheduled sacrifices. The sacrifice times were at 368, 555, 754, 919, 1096 and 1292 days. All the animals were sacrificed or died before the 5th interval in the male AL group and before the 6th interval in the female AL group. Thus, we include the animals in the first four intervals for males and in the first five intervals for females. The tumor types included in the analysis are as follows: 1) mononuclear cell leukemia (MCL), 2) pituitary adenoma/carcinoma (PIT), 3) islet adenoma/carcinoma (Islet), 4) thyroid c-cell adenoma/carcinoma (C-Cell), 5) pheochromocytoma benign/malignant and adrenal mixed malignant (PHEO), 6) mammary fibroadenoma (FIB, AD), females, 7) skin fibroma (FIBR), females, 8) interstitial cell adenoma (INTR), males, 9) clitoral gland adenoma/carcinoma (CLIT), females.

Table 3 compares the results from the CA test, the Poly- $k$  test with several fixed  $k$  values, the generalized Poly- $k$  test based on the estimated value of  $k$  [16] and the proposed test. In general, the  $p$ -values of the proposed test are not always close to those of the generalized Poly- $k$  test. This is because the proposed test is totally nonparametric, while the Weibull distribution of the tumor onset is assumed in the Poly- $k$  test. The effect of calorie restriction appears to be significant for PIT, PHEO, FIB AD, INTR, CLIT and insignificant for C-CELL and FIBR at the significance level of 0.05 from all the tests considered in this study. For MCL and Islet, the results do not agree among the different tests considered in this study. The generalized Poly- $k$  test appears to be more

Table 3:  $p$ -values for the CA test, Poly- $k$  test with fixed  $k$ , the generalized Poly- $k$  test [16] and the proposed method.

Tumor	Sex	CA	Poly- $k$ with fixed $k$			Generalized Poly- $k$	Proposed test
			$k = 1.5$	$k = 3$	$k = 6$		
MCL <sup>a</sup>	M	0.924	0.902	0.812	0.526	0.856	0.681
	F	0.436	0.401	0.165	0.010	0.220	0.041
PIT <sup>b</sup>	M	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$
	F	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$
Islet <sup>c</sup>	M	0.141	0.097	0.057	0.017	0.083	0.030
	F	0.293	0.227	0.157	0.068	0.222	0.039
C-Cell <sup>d</sup>	M	0.337	0.294	0.242	0.156	0.216	0.064
	F	0.294	0.203	0.114	0.028	0.171	0.120
PHEO <sup>e</sup>	M	0.050	0.031	0.017	0.005	0.024	0.001
	F	0.030	0.015	0.006	0.001	0.016	0.001
FIB AD <sup>f</sup>	F	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$
FIBR <sup>g</sup>	F	0.386	0.326	0.255	0.147	0.255	0.222
INTR <sup>h</sup>	M	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$
CLIT <sup>i</sup>	F	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$	$\simeq 0$

<sup>a</sup> mononuclear cell leukemia<sup>b</sup> pituitary adenoma/carcinoma<sup>c</sup> islet adenoma/carcinoma<sup>d</sup> thyroid c-cell adenoma/carcinoma<sup>e</sup> pheochromocytoma benign/malignant and adrenal mixed malignant<sup>f</sup> mammary fibroadenoma, females<sup>g</sup> skin fibroma, females<sup>h</sup> interstitial cell adenoma, males<sup>i</sup> clitoral gland adenoma/carcinoma, females

conservative than the proposed test.

The MLEs of the parameters and  $S(t_j)$  are given in Table 4. We chose the MCL and PIT tumor types for illustration. As discussed in Section 3, the data are pooled across dose groups in this estimation. The monotonicity constraints given in (1) are enforced in both sexes of the PIT group. This constraint is essential for the proposed test because the time-at-risk weight defined in (6) exceeds 1 if the constraint is not enforced. The monotonicity constraint is not violated in the MCL groups. Note that the estimation was not available at the interval of [1097, 1292] for the males due to lack of sacrificed animals.

## 6 DISCUSSION

In this article, we derived explicit solutions for the constrained nonparametric MLEs which do not require the use of computer-intensive numerical methods used in previous work [19, 20]. A modification of the Poly- $k$  test is made by replacing the time-at-risk weight to a function of the tumor onset survival function. The tumor onset survival function can be estimated using the

Table 4: Constrained maximum likelihood estimates for the PCR data with MCL and PIT tumor types.

Group	Interval	$\hat{f}$ (se)	$\hat{g}$ (se)	$\hat{\pi}$ (se)	$\hat{S}(t_j)$ (se)
MCL <sup>a</sup>	0-368	1.000 (0.049)	0.995 ( 0.003)	1.000 (0.196)	1.000 (0.202)
Male	369-555	0.985 (0.008)	0.964 (0.009)	1.000 (0.163)	0.985 (0.165)
	556-754	1.000 (0.065)	0.673 (0.048)	0.478 (0.081)	0.471 (0.079)
	755-919	0.781 (0.083)	0.516 (0.061)	0.476 (0.088)	0.366 (0.069)
	920-1096	1.000 (0.183)	0.188 (0.042)	0.443 (0.070)	0.340 (0.092)
MCL	0-368	1.000 (0.048)	0.991 ( 0.046)	0.995 (0.035)	0.995 (0.059)
Female	369-555	0.995 (0.007)	0.963 (0.009)	1.000 (0.161)	0.995 (0.165)
	556-754	0.970 (0.023)	0.806 (0.026)	0.792 (0.075)	0.764 (0.077)
	755-919	0.864 (0.064)	0.565 (0.048)	0.583 (0.084)	0.487 (0.069)
	920-1096	0.807 (0.128)	0.361 (0.071)	0.583 (0.124)	0.393 (0.088)
	1097-1292	1.000 (0.302)	0.286 (0.062)	0.714 (0.088)	0.299 (0.138)
PIT <sup>b</sup>	0-368	1.000 (0.051)	0.995 (0.004)	0.885 (0.063)	0.885 (0.077)
Male	369-555	0.995 (0.011)	0.960 (0.014)	0.636 (0.100)	0.633 (0.103)
	†556-754	0.887 (0.050)	0.773 (0.044)	0.552 (0.085)	0.487 (0.083)
	†755-919	1.000 (0.083)	0.425 (0.028)	0.552 (0.031)	0.487 (0.078)
	920-1096	1.000 (0.152)	0.197 (0.023)	0.552 (0.039)	0.487 (0.149)
PIT	0-368	1.000 (0.050)	0.990 ( 0.005)	0.963 (0.036)	0.963 (0.060)
Female	369-555	0.998 (0.006)	0.959 (0.011)	0.850 (0.079)	0.848 (0.088)
	556-754	1.000 (0.048)	0.782 (0.039)	0.429 (0.066)	0.428 (0.074)
	†755-919	0.857 (0.11)	0.574 (0.069)	0.430 (0.085)	0.368 (0.079)
	†920-1096	0.602 (0.126)	0.437 (0.079)	0.714 (0.127)	0.368 (0.098)
	1097-1292	1.000 (0.264)	0.286 (0.094)	0.714 (0.142)	0.368 (0.154)

<sup>a</sup> mononuclear cell leukemia<sup>b</sup> pituitary adenoma/carcinoma

†The monotonicity constraints given in (1) are enforced in this interval.

proposed maximum likelihood estimation method. Since our test uses the estimated tumor onset survival function to replace the time-at-risk weight, we can avoid a possible misspecification of  $k$  in the Poly- $k$  test.

The proposed estimators are derived using three survival functions for tumor onset time, death from the tumor, and death from competing risks, while Malani and Van Ryzin [23] or Williams and Portier [22] use tumor incidence rates and death rates with or without tumors. The unconstrained MLEs for the tumor onset survival function in our model are the same as those in [22]. It can be shown that the unconstrained estimate of  $\lambda_m$  in [22] is equivalent to  $1 - (f_m \pi_m / \pi_{m-1})$  in the second case of Theorem 2-1. However, the constraints of the  $f_j$  and the relationship between the two survival functions  $F$  and  $S$  in our model introduce different constrained MLEs. There is no pattern that the constrained estimates by the proposed method are consistently higher or lower than the estimates by the other two approaches.

The proposed test is used to analyze the data from the Project on Calorie Restriction. Our method sometimes gave different test results from the Poly-3 test. This is caused by the difference between the test using a fixed  $k$  and a test using the time-at-risk weight, which does not include  $k$ , based on the proposed MLE of the tumor onset survival function. This makes the proposed test nonparametric, which does not depend on the assumption on the tumor onset distribution.

Most of the recent long-term animal carcinogenicity studies in survival-sacrifice experiments have been designed with a single terminal sacrifice, although interim sacrifices are not uncommon in these types of experiment. In some carcinogenicity studies, it may be useful to sacrifice a few animals in the course of the study for obtaining more information about the tumor incidence rates among live animals. If a goal is to estimate the tumor onset rather than tumor death with low tumor lethality, a study could include interim sacrifices at specified time points to estimate the tumor incidence rate [33].

Besides the PCR study, some examples of experiments with interim sacrifices include the ED<sub>01</sub> study [16, 34] and the experiment on benzidine dihydrochloride [3, 23, 35]. As seen in Section 2, interim sacrifices of a few animals are essential for obtaining information on the tumor status among live animals when tumors cannot be individually classified as incidental or fatal [36]. Extending the proposed method to the analysis of single terminal sacrifice is deferred to a future study.

## REFERENCES

1. Tarone R. Tests for trend in life table analyses. *Biometrika* 1975; **62**:679-682.
2. Peto R. Guidelines on the analysis of tumour rates and death rates in experimental animals. *British Journal of Cancer* 1974; **29**:101-105.
3. Kodell RL, Nelson CJ. An illness-death model for the study of the carcinogenic process using survival/sacrifice data. *Biometrics* 1980; **36**:267-277.
4. Turnbull BW, Mitchell TJ. Nonparametric estimation of the distribution of time to onset for specific diseases in survival/sacrifice experiments. *Biometrics* 1984; **40**:41-50.
5. Archer LE, Ryan LM. On the role of cause-of-death data in the analysis of rodent tumorigenicity experiments. *Applied Statistics* 1989; **38**:81-93.
6. Stallard N, Whitehead A. Modified Weibull multi-state models for the analysis of animal carcinogenicity data. *Environmental and Ecological Statistics* 2000; **7**:117-133.
7. Lagakos SW. An evaluation of some two-sample tests used to analyze animal carcinogenicity experiments. *Utilitas Mathematica* 1982; **21B**:239-260.
8. Lagakos SW, Louis TA. Use of tumour lethality to interpret tumorigenicity experiments lacking cause-of-death data. *Applied Statistics* 1988; **37**:169-179.
9. Bailer AJ, Portier CJ. Effects of treatment-induced mortality and tumor-induced mortality on tests for carcinogenicity in small samples. *Biometrics* 1988; **44**:417-431.
10. Cochran WG. Some methods for strengthening the common  $\chi^2$  tests. *Biometrics* 1954; **10**:417-451.
11. Armitage P. Tests for linear trends in proportions and frequencies. *Biometrics* 1955; **11**:375-386.
12. Bieler GS, Williams RL. Ratio estimates, the delta method, and quantal response tests for increased carcinogenicity. *Biometrics* 1993; **49**:793-801.

13. STP Peto Analysis Working Group. Draft recommendations on classification of rodent neoplasms for Peto analysis. *Toxicologic Pathology* 2001a; **29**:265-268.
14. STP Peto Analysis Working Group. The Society of Toxicologic Pathology's position on statistical methods for rodent carcinogenicity studies. *Toxicologic Pathology* 2001b; **29**:670-672.
15. STP Peto Analysis Working Group. The Society of Toxicologic Pathology's recommendations on rodent carcinogenicity studies. *Toxicologic Pathology* 2002; **30**:415-418.
16. Moon H, Ahn H, Kodell RL, Lee JJ. Estimation of  $k$  for the Poly- $k$  test with application to animal carcinogenicity studies. *Statistics in Medicine* 2003; **22**:2619-2636.
17. Kodell RL, Ahn H. Age-adjusted trend test for the tumor incidence rate. *Biometrics* 1997; **53**:1467-1474.
18. Moon H, Ahn H, Kodell, RL. An age-adjusted bootstrap-based Poly- $k$  test. *Statistics in Medicine* 2005; **24**:1233-1244.
19. Ahn H, Moon H, Kodell RL. Attribution of tumor lethality and estimation of time to onset of occult tumors in the absence of cause-of-death information. *Applied Statistics* 2000; **49**:157-169.
20. Moon H, Ahn H, Kodell RL. Extension of Peto's test by attribution of tumor lethality in the absence of cause-of-death information. *Biometrical Journal* 2002; **44**:982-1001.
21. Williams PL, Portier CJ. Analytic expressions for maximum likelihood estimators in a non-parametric model of tumor incidence and death. *Communications in Statistics - Theory and Methods* 1992a; **21**:711-732.
22. Williams PL, Portier CJ. Explicit solutions for constrained maximum likelihood estimators in survival/sacrifice experiments. *Biometrika* 1992b; **79**:717-729.
23. Malani HM, Van Ryzin J. Comparison of two treatments in animal carcinogenicity experiments. *Journal of the American Statistical Association* 1988; **83**:1171-1177.

24. Kalbfleisch JD, Prentice RL. *The Statistical Analysis of Failure Time Data*. Wiley: New York, 1980; pp14-15.
25. Portier CJ, Hedges J, Hoel DG. Age-specific models of mortality and tumor onset for historical control animals in the National Toxicology Program's carcinogenicity experiments. *Cancer Research* 1986; **46**:4372-4378.
26. U.S. FDA. Guidance for industry: Statistical aspects of the design, analysis, and interpretation of chronic rodent carcinogenicity studies of pharmaceuticals. *Federal Register* 2001; **66(89)**:23266-23267.
27. Mancuso, J. Y., Ahn, H, Chen, J. J. and Mancuso, J. P. Age-adjusted exact trend tests in the event of rare occurrences. *Biometrics* 2002; **58**:403-412.
28. Pariza MW, Boutwell RK. Historical perspective: Calories and energy expenditure in carcinogenesis. *The American Journal of Clinical Nutrition* 1987; **45** (Suppl.1):151-156.
29. Roe FJC. Non-genotoxic carcinogenesis: Implications for testing and extrapolation to man. *Mutagenesis* 19889; **4**:407-411.
30. Roe FJ, Lee PN, Conybeare G, Tobin G, Kelly D, Prentice D, Matter B. Risks of premature death and cancer predicted by body weight in early adult life. *Human and Experimental Toxicology* 1991; **10**:285-288.
31. Hart R, Neumann D, Robertson R. Dietary restriction: *Implications for the design and interpretation of toxicity and carcinogenicity studies*. ILSI Press, Washington, DC.
32. Thurman JD, Bucci TJ, Hart RW, Turturro A. Survival, body weight, and spontaneous neoplasms in *ad libitum*-fed and food-restricted Fischer-344 rats. *Toxicologic Pathology* 1994; **22**:1-9.
33. Fairweather WR, Bhattacharyya AB, Ceuppens PP, Heimann G, Hothorn LA, Kodell RL, Lin KK, Mager H, Middleton BJ, Slob W, Soper KA, Stallard N, Ventre J, Wright J. Biostatistical methodology in carcinogenicity studies. *Drug Information Journal* 1998; **32**:401-421.

34. Littlefield NA, Farmer JH, Gaylor DW, Sheldon WG. Effects of dose and time in a long-term, low-dose carcinogenic study. *Journal of Environmental Pathology and Toxicology* 1980; **3**:17-34.
35. Ahn H, Kodell RL. Estimation and testing of tumor incidence rates in experiments lacking cause-of-death data. *Biometrical Journal* 1995; **37**:745-763.
36. McKnight B, Crowley J. Tests for differences in tumor incidence based on animal carcinogenesis experiments. *Journal of the American Statistical Association* 1984; **79**:639-648.

## APPENDIX: Proof of the theorem

In order to prove the theorem, we introduce the following Lemma:

### Lemma 1

$$\frac{N_j^* \pi_j}{N_j^* \pi_j + b_{1j}} \leq g_j^* \iff \frac{b_{1j}}{ad_j + b_{1j}} \geq \pi_j \iff \frac{N_j^*}{N_{j-1}} \leq g_j^*,$$

where

$$g_j^* = \frac{1}{1 - \pi_j} \left( \frac{N_j^* + b_{1j}}{N_{j-1}} - \pi_j \right).$$

### Proof of Lemma 1

$$\begin{aligned} \frac{N_j^* \pi_j}{N_j^* \pi_j + b_{1j}} \leq \frac{1}{1 - \pi_j} \left( \frac{N_j^* + b_{1j}}{N_{j-1}} - \pi_j \right) &\iff \left\{ N_j^* + b_{1j} - \frac{N_j^*(N_j^* + b_{1j})}{N_{j-1}} \right\} \pi_j \leq \frac{b_{1j}(N_j^* + b_{1j})}{N_{j-1}} \\ &\iff \pi_j \leq \frac{b_{1j}}{ad_j + b_{1j}} \end{aligned}$$

$$\begin{aligned} \frac{N_j^*}{N_{j-1}} \leq \frac{1}{1 - \pi_j} \left( \frac{N_j^* + b_{1j}}{N_{j-1}} - \pi_j \right) &\iff N_j^*(1 - \pi_j) \leq N_j^* + b_{1j} - N_{j-1}\pi_j \\ &\iff \pi_j \leq \frac{b_{1j}}{ad_j + b_{1j}} \quad \square \end{aligned}$$

**Proof of Theorem 2.1** First we fix  $\pi_j \in [0, 1]$ . Suppose that  $g_j \in [0, 1]$  is also fixed. Solving (2), the log-likelihood function is maximized when

$$f_j = f_j(g_j, \pi_j) = \begin{cases} \frac{N_j^* + b_{1j}}{N_{j-1}\{g_j + (1-g_j)\pi_j\}} & \text{if } g_j \geq g_j^* \\ 1 & \text{if } g_j \leq g_j^*. \end{cases}$$

Substituting the above  $f_j = f_j(g_j, \pi_j)$  into (3) with applying Lemma 1,

$$g_j = g_j(\pi_j) = \begin{cases} \frac{N_j^* \pi_j}{N_j^* \pi_j + b_{1j}} & \text{if } \pi_j \geq \frac{b_{1j}}{ad_j + b_{1j}} \\ \frac{N_j^*}{N_j^* + ad_j + b_{1j}} & \text{if } \pi_j \leq \frac{b_{1j}}{ad_j + b_{1j}}. \end{cases}$$

Therefore, the maximization can be done with a continuous and piecewise differentiable function

$$\tilde{l}_j(\pi_j) = l_j(f_j(g_j(\pi_j), \pi_j), g_j(\pi_j), \pi_j) = \begin{cases} \tilde{l}_j^{(1)} & \text{if } \pi_j \geq \frac{b_{1j}}{ad_j + b_{1j}} \\ \tilde{l}_j^{(2)} & \text{if } \pi_j \leq \frac{b_{1j}}{ad_j + b_{1j}}, \end{cases}$$

where  $l = \sum_{j=1}^s l_j(f_j, g_j, \pi_j) + C$ . Note that

$$\min \left\{ \frac{b_{2j}}{a_{2j} + b_{2j}}, \frac{b_{1j}}{ad_j + b_{1j}} \right\} \leq \frac{b_{1j} + b_{2j}}{ad_j + b_{1j} + a_{2j} + b_{2j}} \leq \max \left\{ \frac{b_{2j}}{a_{2j} + b_{2j}}, \frac{b_{1j}}{ad_j + b_{1j}} \right\}.$$

From (4) with substitutions,  $\tilde{l}_j^{(1)}$  has only one critical point at  $\pi_j = b_{2j}/(a_{2j} + b_{2j})$  which is a maximum point and  $\tilde{l}_j^{(2)}$  has only one critical point at  $\pi_j = (b_{1j} + b_{2j})/(ad_j + b_{1j} + a_{2j} + b_{2j})$  which is a maximum point. Therefore, if  $b_{2j}/(a_{2j} + b_{2j}) \geq b_{1j}/(ad_j + b_{1j})$ , then  $\tilde{l}_j$  is increasing in  $\pi_j$  on  $(0, b_{2j}/(a_{2j} + b_{2j}))$  and decreasing on  $(b_{2j}/(a_{2j} + b_{2j}), 1)$ . If  $b_{2j}/(a_{2j} + b_{2j}) \leq b_{1j}/(ad_j + b_{1j})$ , then  $\tilde{l}_j$  is increasing in  $\pi_j$  on  $(0, (b_{1j} + b_{2j})/(ad_j + b_{1j} + a_{2j} + b_{2j}))$  and decreasing on  $((b_{1j} + b_{2j})/(ad_j + b_{1j} + a_{2j} + b_{2j}), 1)$ . We maximize  $\tilde{l}_j$  for  $\pi_j \in [\pi_{j+1}f_{j+1}, 1]$ , then the result is obtained.  $\square$