

An Alternative to the Peto Model for 2-Year Carcinogenicity Studies and to Tarone's Test for Trend in Censored Survival Data

Arthur J. Roth, Ph.D.

Midwest Biopharmaceutical Statistics Workshop

Muncie, IN

May 19, 2009

OUTLINE

1. Statistical Framework of Problem
 - A. Survival data basics: brief review
 - B. Carcinogenicity studies
 - C. Current “standard” analyses
 - D. Drawbacks
2. Proposal for new improved method
3. Simulations
4. Conclusions and ongoing work

SIMPLE (NONPARAMETRIC) SURVIVAL DATA EXAMPLE

(A) 8, 20+, 20+ (B) 12, 20+, 20+ (C) 6, 12, 20+ (D) 5, 7+, 12

“+” = censored → non-censored event times are 5, 6, 8, 12

Time 5

0	0	0	1

1
11

3 3 3 3

Add first row across tables

$$\implies (1, 1, 2, 2) = (O_1, O_2, O_3, O_4)$$

Time 6

0	0	1	0

1
10

3 3 3 2

is total number of uncensored events
in each group

Time 8

1	0	0	0

1
8

3 3 2 1

Time 12

0	1	1	1

3
5

2 3 2 1

CRUCIAL DISTRIBUTIONAL PROPERTIES

(1) View **red** as fixed (2) View **teal** as random (3) Tables \approx independent

Time 5

X_1	X_2	X_3	X_4	1
				11

3 3 3 3

$(X_1, X_2, X_3, X_4)^T$ has known mean \mathbf{E}_5 and covariance matrix \mathbf{T}_5 (under H_0)

Time 6

X_1	X_2	X_3	X_4	1
				10

3 3 3 2

$(X_1, X_2, X_3, X_4)^T$ has known mean \mathbf{E}_6 and covariance matrix \mathbf{T}_6 (under H_0)

(similarly for \mathbf{E}_8 , \mathbf{T}_8 , \mathbf{E}_{12} , and \mathbf{T}_{12})

Time 8

X_1	X_2	X_3	X_4	1
				8

3 3 2 1

Time 12

X_1	X_2	X_3	X_4	3
				5

2 3 2 1

$\mathbf{O} = (O_1, O_2, O_3, O_4)^T = (\sum X_1, \sum X_2, \sum X_3, \sum X_4)^T$ has (under H_0) known mean $\mathbf{E} = \sum \mathbf{E}_t$ and Covariance matrix $\mathbf{T} = \sum \mathbf{T}_t$ (convolution under independence). It is also known to be asymptotically multivariate normal (with a singularity). Exact small sample permutation distribution is also known.

STANDARD (NONPARAMETRIC) SURVIVAL ANALYSIS

$k = \#$ groups (no limit)

$\mathbf{O} = (O_1 \ O_2 \ \dots \ O_k)^T \sim$ mean \mathbf{E} , covariance matrix \mathbf{T}

May want to test for either

homogeneity (unordered groups)

trend (ordered groups, e.g., C L M H)

Homogeneity (logrank test)

$$(\mathbf{O} - \mathbf{E})^T \mathbf{T}^{-1} (\mathbf{O} - \mathbf{E}) \sim X^2(k - 1) \quad (\text{generalized inverse})$$

Trend (Tarone's test)

(a) choose a monotone contrast, e.g., $\mathbf{c} = (-3 \ -1 \ 1 \ 3)^T$

(b) $\mathbf{c}^T (\mathbf{O} - \mathbf{E}) / \text{sqrt}(\mathbf{c}^T \mathbf{T} \mathbf{c}) \sim N(0,1)$

1-sided (in either direction) or 2-sided

CARCINOGENICITY STUDY (DESIGN)

- Animals (rats or mice) dosed daily for 2 years (full lives)
- Groups always ordered (C L M H or C L M H VH)
 - NOTE: Control usually vehicle for drug. May also have second type of control, usually (but not always!) positive control. If so, do two analyses:
 - (a) compare the controls
 - (b) control vs. dose groups (control prespecified, usually vehicle)
- Main parameters: tumor incidence (more precisely, time to tumor onset) for each tumor type (≈ 50) separately.
 - NOTE: Survival time itself is also of (secondary) interest
- $k = \#$ groups (per analysis) is almost always
 - 2 (compare controls) or
 - 4 or 5 (controls + 3 or 4 doses of test compound)
- (controversial) Pathologist classifies each occurrence of a tumor as either fatal (contributed to animal's death) or incidental (animal died of other causes)

CARCINOGENICITY STUDY (ANALYSIS)

- Tumor risk \uparrow with age \rightarrow elements of survival analysis
- Trend tests
 - (a) FDA requires for tumor incidence (1-sided)
 - (b) usually also done for survival analysis (2-sided)
- For tumor incidence, two types of 2xk tables
 - (a) fatal tumors: same as survival analysis, where event = death from tumor type of interest (and all other deaths are censored observations)
 - (b) incidental tumors: divide time line into prespecified intervals (e.g., 0-12 mo., 12-18, 18-21, 21-24, final sac). Form 2xk table within each interval, where (in each grp)
 - (i) numerator = # incidental tumors (type of interest)
 - (ii) denominator = # died in interval
- Structure of data identical to survival data!!!!

WHERE ARE WE?

- Same tests are appropriate for survival and carcinogenicity data (same statistical data structure)

NOTE: The names logrank test and Tarone's test are used only for survival. The trend test for carcinogenicity is called the Peto test.

- We are interested only in trend tests on this data structure and (surprisingly important) only when $2 \leq k \leq 5$.
- SIDE NOTE: Data structure valid for carcinogenicity studies only due to controversial assumption. Some statisticians advocate replacing the Peto test by Poly-3 trend test, which doesn't require this information from the pathologist, but makes another challengeable assumption instead. (Will briefly describe.) But even the Poly-3 test is contrast-based.

DRAWBACKS TO CONTRAST-BASED TREND TESTS

- The choice of the contrast is arbitrary, and different contrasts can produce very different results.

NOTE: Commonly used contrasts are equally spaced (-3 -1 1 3) or corresponding to actual doses, perhaps (0 10 30 100), or logs of actual doses (leaving controls at 0).

- More important: No contrast truly reflects the hypothesis of interest. In particular, negative coefficients mean that tumors at certain doses actually reduce significance!

EXAMPLES

(-3 -1 1 3): Low-dose tumors help the test compound
(0 10 30 100) → (-7 -5 -1 13) even mid-dose tumors help
(0 10 30 400) → (-11 -10 -8 29) even worse

So tumor pattern (2 10 8 7) would rarely be significant for any of these contrasts.

IMPORTANT: Any monotone contrast optimal against some alternative

DEVELOPMENT OF PROPOSED NEW TEST

- Formulate hypotheses in terms of “right” r.v.’s such that
 H_0 : all groups have equal tumor incidence rates
 H_A : tumor incidence is a monotone increasing function of dose (in some sense)
- Problem in isotonic regression with known solution (LRT)
 $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ known
 $H_0: \mu_1 = \mu_2 = \dots = \mu_k$
 $H_A: \mu_1 \leq \mu_2 \leq \dots \leq \mu_k$ with at least one “<”
- For our data, how well does this fit?
 \mathbf{O} is only asymptotically normal (ok, exact test if small N’s)
 \mathbf{E} and \mathbf{T} known only under H_0 (test valid, won’t be LRT)
 $\mathbf{E} = (E_1, E_2, \dots, E_k)$ known but NOT equal under H_0 or monotone under H_A (unequal Ns and survival)
- SOLUTION: $R_i = O_i / E_i$ (check out above properties)
Means intuitively monotone under H_A because E_i ’s already take N’s and survival into account --- turns out to be only approximately true (but good enough)

HYPOTHESES EXPRESSED IN TERMS OF R_i 's

- $R_i = O_i / E_i$ and $\mathbf{R} = (R_1, R_2, \dots, R_k)^T$
- Under H_0 , $E(R_i) = 1$ and $E(\mathbf{R}) = \mathbf{1} = (1, 1, \dots, 1)^T$
- $\Psi = \text{Cov}(\mathbf{R})$ [has a singularity, just like \mathbf{T}]
- Under H_0 , Ψ known with elements $\psi_{ij} = \text{Cov}(R_i, R_j) = \tau_{ij} / (E_i E_j)$
- $H_0: E(R_1) = E(R_2) = \dots = E(R_k)$ [don't care that common value=1]
 $H_A: E(R_1) \leq E(R_2) \leq \dots \leq E(R_k)$ with at least one "<"

Fits problem with known solution --- what is the solution?

SOLUTION TO ISOTONIC REGRESSION PROBLEM

STEP 1: Transform variables

$$D_i = R_{i+1} - R_i \quad (i=1, 2, \dots, k-1), \quad \mathbf{D} = (D_1 \dots D_{k-1})^T$$

IMMEDIATELY IMPORTANT

(a) Gets rid of the singularity

(b) Reduces dimension by 1 (must solve in ≤ 4 dimensions)

$$H_0: E(D_i) = 0 \text{ for all } i$$

$$H_A: E(D_i) \geq 0 \text{ for all } i \text{ with at least one } ">"$$

Note that $\mathbf{D} = \mathbf{P}\mathbf{R}$, where \mathbf{P} is the following $(k-1) \times k$ matrix:

$$\begin{array}{cccccc} -1 & 1 & 0 & \dots & 0 & \\ 0 & -1 & 1 & 0 & \dots & 0 \\ \vdots & & & & & \\ 0 & \dots & 0 & -1 & 1 & \end{array}$$

$$\mathbf{\Delta} = \text{Cov}(\mathbf{D}) = \mathbf{P} \text{Cov}(\mathbf{R}) \mathbf{P}^T = \mathbf{P} \mathbf{\Psi} \mathbf{P}^T \text{ (known)}$$

Note that the parameter space is restricted to $E(D_i) \geq 0$
under both H_0 and H_A (so MLE must also be restricted)

SOLUTION TO ISOTONIC REGRESSION PROBLEM

STEP 2: restricted “MLE” of $E(\mathbf{D})$

Let $\mathbf{D}^* = (D_1^* \dots D_{k-1}^*)^T$ be the “MLE” of $E(\mathbf{D})$
restricted to $D_i^* \geq 0$ for all i

In other words, \mathbf{D}^* lies in Q , the nonnegative orthant in $(k-1)$ -space.

(“MLE” in quotes because $\text{Cov}(\mathbf{D}) = \mathbf{\Delta}$ only under H_0)

- (a) Obviously, if all $D_i \geq 0$ (i.e., \mathbf{D} lies in Q), then $\mathbf{D}^* = \mathbf{D}$
- (b) Otherwise, algorithm for finding \mathbf{D}^* is interesting, but no time to go through it

Some intuition

\mathbf{D}^* is the projection of \mathbf{D} on Q using the distance measure defined by $\mathbf{\Delta}^{-1}$, i.e.,

\mathbf{D}^* is the point in Q which minimizes $(\mathbf{D} - \mathbf{D}^*)^T \mathbf{\Delta}^{-1} (\mathbf{D} - \mathbf{D}^*)$

(motivate using Euclidean distance and zeros)

SOLUTION TO ISOTONIC REGRESSION PROBLEM

STEP 3: Test statistic

Reject H_0 for large values of

$T = (\mathbf{D}^*)^T \mathbf{\Delta}^{-1} \mathbf{D}^*$, which is the length of \mathbf{D}^* using the same distance measure used when finding \mathbf{D}^* to begin with
(large values of \mathbf{D}^* make intuitive sense)

STEP 4: Null distribution

(a) REMINDER: For small numbers of events, use exact permutation distribution of T , and we're done

(b) Asymptotically, it's a mixture of chi-square distributions:

$$P(T \geq t) = \sum_{(0 \leq j \leq k-1)} Q(j, k-1) P(X_j^2 \geq t) \quad [\text{under } H_0]$$

where X_j^2 is a chi-square(j) r.v. ($X_0^2 \equiv 0$) and

$Q(j, k-1)$ is the probability that \mathbf{D}^* has exactly j non-zero elements

(It remains only to find the $Q(j, k-1)$'s.)

SOLUTION TO ISOTONIC REGRESSION PROBLEM

STEP 5: Finding the mixing probabilities $Q(j,k-1)$

NOTE: These mixing probabilities depend on Δ . Let ρ_{ij} and $\rho_{ij.k}$ represent the respective correlations and partial correlations corresponding to the covariances Δ .

(A) For $k-1 \geq 5$ ($k \geq 6$), there are no known formulas, only approximations that are sometimes fairly inaccurate \rightarrow it's important that $k \leq 5$ for carc studies, though not an issue for the exact permutation test

(B) There are known closed-form expressions for $Q(j,k-1)$ for $k-1 = 1,2,3$ ($k=2,3,4$). Here they are for $k-1=3$.

$$Q(0,3) = 1/2 - [\text{Arccos}(\rho_{12}) + \text{Arccos}(\rho_{13}) + \text{Arccos}(\rho_{23})] / (4\pi)$$

$$Q(1,3) = 3/4 - [\text{Arccos}(\rho_{12.3}) + \text{Arccos}(\rho_{13.2}) + \text{Arccos}(\rho_{23.1})] / (4\pi)$$

$$Q(2,3) = 1/2 - Q(0,3) \quad Q(3,3) = 1/2 - Q(1,3)$$

(C) For $k-1=4$ ($k=5$)

(i) $Q(j,4)$ can be calculated for $j=0,1,2,3$ from the $Q(j,3)$'s using recursion formulas.

(ii) There is an integral (not closed form) expression for $Q(4,4)$ in the special case of independence of the original variables (so only adjacent D_i 's are correlated --- NOT the case in our application).

(iii) Abrahamson (1964) gives an algorithm (with typos) that expresses $Q(4,4)$ as a linear combination of six terms involving the special case (ii). **DONE!!!!!!**

SIMULATIONS

- Four groups only (after all that work, no use for Q(4,4))
- N=60 animals per group
- 10000 replications for null cases, 4000 for non-null
- Compared new procedure against scores of
 $(-3 \ -1 \ 1 \ 3)$
 $(0 \ 10 \ 30 \ 100) = (-7 \ -5 \ -1 \ 13)$
 $(0 \ 10 \ 30 \ 400) = (-11 \ -10 \ -8 \ 29)$
- Exact tests (all procedures) if ≤ 10 tumor occurrences
- Parameters allowed to vary
 - (a) For each of C L M H, specify % surviving death from competing risks (not due to the tumor type of interest)
 - (b) For each of C L M H, specify % with tumor onset before 2 years if competing risks are ignored
 - (c) Parameter that controls what % of tumors are fatal

SIMULATIONS --- MODEL

0---(t₁)----T₁----(t₂)----T₁+T₂

0------(t₃)-----T₃

T₁=time to tumor onset

T₂=time from tumor onset until death from tumor

T₃=time to death from competing risks

Animal dies at $\text{Min}(T_1+T_2, T_3)$ if <2 years (scaled s.t. $t=1$)

- If $T_3 < T_1$, dies tumor free
- Otherwise if $T_3 < T_1 + T_2$, dies with incidental tumor
- Otherwise dies with fatal tumor

Animal survives to final sacrifice if $\text{Min}(T_1+T_2, T_3) \geq 2$ years

- If $T_1 \leq 2$ years, incidental tumor
- If $T_1 > 2$ years, tumor free

SIMULATIONS --- PROBABILITY DISTRIBUTIONS

T₁=time to tumor onset

Weibull distribution with cubic decay:

$$F(t) = 1 - \exp\{-\beta t^3\}$$

Each group's β chosen separately so that

$F(1)$ = that group's tumor rate (ignoring competing risks)

T₂=time from tumor onset until death from tumor

Modeled as a constant (unrealistic)

values used: 30 (~90% fatal), 200 (30-40%), 350 (~10%)

T₃=time to death from competing risks

Modified Weibull with (mainly) 7th power decay

$$F(t) = 1 - \exp\{-\varphi (.025t + \beta t^7)\}, \text{ with } \varphi=1 \text{ for controls}$$

Find β based on $F(1)$ for controls, then find φ for each group based on that group's $F(1)$

SIMULATIONS --- WHAT CASES?

T_1 =time to tumor onset

Control background rates of 2%, 5%, 10%, and 20%

Treated: same as control (10000)

various increasing patterns (a few not monotone) (4000)

T_3 =time to death from competing risks

Equal survival (33% and 50%)

High dose survives longer than rest

High dose dies faster than rest

Turned out: T_2 and T_3 had no bearing on relative performance of the procedures. All the variation was in T_1 (background rate and response pattern).

Will show results only for $T_2=200$ (middle value) and survival of 33% in every group.

Background = 0.02

Survival = .33 in all groups, T2 = 200 (30-40% tumors fatal)

<u>% tumors (C L M H)</u>	<u>new</u>	<u>-3 -1 1 3</u>	<u>0 10 30 100</u> <u>-7 -5 -1 13</u>	<u>0 10 30 400</u> <u>-11 -10 -8 29</u>
2 2 2 2	2.8	2.0	2.7	2.7
2 2 2 10	62	53	65	65
2 2 10 10	66	63	50	42
2 2 6 10	60	58	57	55
2 10 10 10	40	33	22	17
2 6 6 10	41	42	41	38
2 4.5 7.5 10	47	49	43	40
2 6 10 10	48	49	35	29
2 2 15 2	47	9	0.4	0.2
2 15 2 2	3	0.0	0.0	0.0
2 15 15 2	34	0.3	0.0	0.0
2 15 12.5 10	52	23	9	7

Background = 0.05

Survival = .33 in all groups, T2 = 200 (30-40% tumors fatal)

<u>% tumors (C L M H)</u>	<u>new</u>	<u>-3 -1 1 3</u>	<u>0 10 30 100</u> <u>-7 -5 -1 13</u>	<u>0 10 30 400</u> <u>-11 -10 -8 29</u>
5 5 5 5	4.9	4.3	5.1	5.0
5 5 5 15	59	53	65	66
5 5 15 15	64	65	52	40
5 5 10 15	57	59	59	54
5 15 15 15	43	37	22	16
5 10 10 15	42	43	42	39
5 8.3 11.7 15	45	50	44	38
5 10 15 15	48	50	36	27
5 5 20 5	40	14	1	0.3
5 25 5 5	14	0.1	0.0	0.0
5 20 20 5	32	1	0.1	0.0
5 17 15 13	37	21	10	7

CONCLUSIONS AND ONGOING WORK

- New procedure avoids arbitrary choice of contrast
- New procedure clearly superior for background rates of 2% (and presumably anything lower) --- never loses more than trivial amounts of power in any case and beats each of the others by quite a bit in some cases
- For background rates of 5% or more, new procedure in some cases has more than trivially less power than other procedures, but on an overall basis it is (a) clearly better than dose scores and (b) competes at least on a par with equally spaces scores (in my opinion preferable)
- Dose group scores sometimes exceed nominal Type I error rate.
- Some robustness (compared to none at all) against non-monotone effects
- H_0 and H_A reflect true goal of analysis --- not true for contrast-based methods
- Can't handle (rare) study with 6 or more groups for tumor types with too many occurrences to do exact permutation version
- Currently working on simulations that (misapply) the method to the Poly-3 test