Published in final edited form as:

J Biopharm Stat. 2009 September ; 19(5): 755–762. doi:10.1080/10543400903105158.

Assessing Cytotoxic Treatment Effects in Preclinical Tumor Xenograft Models

Jianrong Wu¹ and Peter J. Houghton²

¹ Department of Biostatistics, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105-3678, USA

² Department of Molecular Pharmacology, St. Jude Children's Research Hospital, 262 Danny Thomas Place, Memphis, TN 38105-3678, USA

Abstract

In preclinical solid tumor xenograft experiments, tumor response to cytotoxic agents is often assessed by tumor cell kill. Log10 cell kill (LCK) is commonly used to quantify the tumor cell kill in such experiments. For comparisons of antitumor activity between tumor lines, the LCK values are converted to an arbitrary rating; for example, the treatment effect is considered significant if the $LCK > 0.7$ (Corbett, et al., 2003). The drawback of using such a predefined cutoff point is that it does not account for the true variation of the experiments. In this paper, a nonparametric bootstrap percentile interval of the LCK is proposed. The cytotoxic treatment effect can be assessed by the confidence limits of the LCK. Monte Carlo simulations are conducted to study the coverage probabilities of the proposed interval for small samples. Tumor xenograft data from a real experiment are analyzed to illustrate the proposed method.

Keywords

Bootstrap; Confidence interval; Log10 cell kill; Right censoring; Xenograft

1. Introduction

In cancer drug development, demonstrating anticancer activity in preclinical animal models is important. The National Cancer Institute has promoted a cancer drug screening program since the mid-1950s. The Pediatric Preclinical Testing Program recently conducted a pediatric cancer drug screening program for both *in vitro* and *in vivo* models (Houghton et al., 2007). In *in vivo* testing, human cancer cells from standard tumor lines are engrafted into mice to produce xenograft models. Tumor-bearing mice are randomized into control (*C*) and treatment (*T*) groups, and the maximum tolerated doses of the cytotoxic agents are administered. The volume of each tumor is measured at the initiation of the study and weekly throughout the study period. Mice are euthanized when the tumor volume reaches four times its initial volume, thus resulting in incomplete longitudinal tumor volume data. In such experiments, tumor response to the cytotoxic agents is often assessed by tumor cell kill. Log10 cell kill (LCK) is commonly used to quantify tumor cell kill. Under some assumptions, the LCK is evaluated as a scaled difference in the median of tumor quadrupling times for treatment and control, divided by a multiple of the median tumor doubling time for the control (Corbett, et al., 2003). Tumor quadrupling times of the experimental mice are often subject to right-censoring because of the death of the experimental mice, limitations of the follow-up period, or the number of cured tumors. The censoring of tumor quadrupling time makes it difficult to analyze such preclinical tumor xenograft data. Stuschke et al. (1990) demonstrated that survival analysis should be used to analyze such data. Survival distributions of tumor quadrupling times should be estimated by

using the Kaplan-Meier method and compared by log-rank tests (Cox and Oakes, 1984). The median tumor quadrupling time should be estimated from the Kaplan-Meier survival distribution. For comparisons of antitumor activity between tumor lines, the LCK values are converted to an arbitrary rating. For example, the treatment effect is considered significant if the $LCK > 0.7$ (Corbett, et al., 2003). The drawback of such a predefined cut point is that it does not account for the true variation of the experiments. To overcome it, a nonparametric bootstrap procedure is proposed to provide a confidence interval of the LCK. Monte Carlo simulations are conducted to study the coverage probability of the proposed interval for small samples. Tumor xenograft data from a real experiment are analyzed to illustrate the proposed method.

2. Tumor Quadrupling Time

In *in vivo* solid tumor xenograft experiments, the volume of each tumor is measured on a weekly schedule for logistical reasons. Therefore, the exact time of quadrupling of a tumor is not measured. A naive approach using the last day of observation without tumor quadrupling is biased. The following interpolation formula can be used to calculate the tumor quadrupling time (day),

$$t_e = t_1 + (t_2 - t_1) \log(V_e/V_{t_1}) / \log(V_{t_2}/V_{t_1}),$$

where t_e is the interpolated quadrupling time, t_1 and t_2 are the lower and upper observation times bracketing the quadrupling tumor volume $V_e = 4V_0$, where V_0 is the initial tumor volume. The tumor quadrupling time calculated from the interpolation formula is exact if tumor growth follows an exponential curve and gives good approximations for other tumor growth curves (Wu, 2007).

Let $t_1 < t_2 < \dots < t_k$ be the unique ordered tumor quadrupling times of a study group, d_j be the number of tumor quadruplings that occur at t_j , and n_j the number of individuals at risk immediately prior to t_j . Then the Kaplan-Meier estimator (Cox and Oakes, 1984) of the survival function is

$$\widehat{S}(t) = \prod_{t_j \leq t} \left(1 - \frac{d_j}{n_j}\right).$$

The median tumor quadrupling time can be estimated as

$$\widehat{\tau}(50) = \min \{t_j, \widehat{S}(t_j) < 0.5\}. \quad (1)$$

When the estimated survival function $\widehat{S}(t)$ is exactly equal to 0.5 on interval $[t_j, t_{j+1})$, the median is taken to be $\widehat{\tau}(50) = (t_j + t_{j+1})/2$ (Collett, 2003).

3. Log10 Cell Kill

In the literature, the LCK is quantified as a scaled difference in the median of tumor quadrupling times for treatment and control, divided by a multiple of the median tumor doubling time for the control group. We illustrate this quantification with following assumptions: (a) tumor growth of the control follows an exponential growth curve, (b) mass (and corresponding volume) of a tumor is directly proportional to the number of malignant cells in the mass, and

(c) the treated tumor regrowth curve approximates that of the tumor in untreated controls. Under assumption (a), the volume (V) of a tumor against time (t) satisfies the following equation

$$V(t) = V_0 e^{bt},$$

where V_0 is the initial volume and b is a parameter characterizing the tumor growth rate. Let T and C be the tumor quadrupling times for treatment and control, respectively; then under assumptions (b) and (c), the tumor cell survival fraction of the treatment vs. control groups is given by the ratio

$$S = V(C) / V(T).$$

The LCK is then the negative log10 cell survival fraction; that is,

$$\begin{aligned} \text{LCK} &= -\log_{10}(1/S) \\ &= (T - C) / (\log_2 10 t_D) \\ &\approx \text{TGD} / (3.32 t_D), \end{aligned}$$

where $\text{TGD} = T - C$ is the tumor growth delay and $t_D = \ln 2 / b$ is the tumor doubling time of the control.

Now let $\widehat{t}_c(50)$ and $\widehat{t}_T(50)$ be the estimates of median tumor quadrupling times of the control and treatment groups defined by (1) and $\widehat{t}_D(50)$ be the estimate of the median tumor doubling time of the control, which can be estimated by the slope from fitting the log tumor volume regression line or estimated using the median tumor doubling time of the control via interpolation. Then an estimate of the LCK is given by

$$\widehat{\text{LCK}} = \widehat{\text{TGD}} / \{3.32 \widehat{t}_D(50)\},$$

where $\widehat{\text{TGD}} = \widehat{t}_T(50) - \widehat{t}_c(50)$ is the estimate of TGD.

4. Bootstrap Percentile Interval

The confidence interval is a formal statistical approach used to assess treatment effect. The confidence limits account for both effect size and experimental variation. Therefore, the antitumor activity of a cytotoxic agent can be assessed on the basis of the confidence limits, and no arbitrary cutoff-point is needed.

To construct the LCK confidence interval, let the observed tumor quadrupling times consist of n pairs $(t_{1g}, \delta_{1g}), \dots, (t_{ng}, \delta_{ng})$ of group $g = T, C$, where $t_{ig} = \min(t_{ig}^0, x_{ig})$ is an observed tumor quadrupling time, t_{ig}^0 or an observed censoring time, x_{ig} , and $\delta_{ig} = I(t_{ig} = t_{ig}^0)$ is an indicator of tumor quadrupling, and let the tumor doubling times consist of n observations $\{t_{1D}, \dots, t_{nD}\}$. The bootstrap percentile interval of LCK can be obtained by using the following bootstrap procedures for the right-censored data (Efron, 1981).

1. Independently draw a large number of bootstrap samples,

$\{(t_{1g}^{*b}, \delta_{1g}^{*b}), \dots, (t_{ng}^{*b}, \delta_{ng}^{*b}), b=1, \dots, M\}$, with each bootstrap sample obtained by sampling with replacement from $\{(t_{1g}, \delta_{1g}), \dots, (t_{ng}, \delta_{ng})\}$, $g = T, C$.

2. Independently draw a large number of bootstrap samples, $\{t_{1D}^{*b}, \dots, t_{nD}^{*b}, b=1, \dots, M\}$, with each bootstrap sample obtained by sampling with replacement from $\{t_{1D}, \dots, t_{nD}\}$.
3. Calculate median estimates of Kaplan-Meier curves for bootstrap samples obtained in steps 1) and 2) above; for example, $\widehat{t}_g^{*b}(50)$ and $\widehat{t}_d^{*b}(50)$, $g = T, C$ and $b = 1, \dots, M$.
4. Calculate the LCK value for each bootstrap sample,

$$\widehat{LCK}^{*b} = \{\widehat{t}_T^{*b}(50) - \widehat{t}_C^{*b}(50)\} / \{3.32\widehat{t}_D^{*b}(50)\}, b=1, \dots, M.$$
5. The bootstrap standard error estimate of the $\widehat{LCK}$ is given by

$$\widehat{se}(\widehat{LCK}) = \left\{ \sum_{b=1}^M (\widehat{LCK}^{*b} - \widehat{\overline{LCK}}^*)^2 / M \right\}^{1/2}$$

$$\text{where } \widehat{\overline{LCK}}^* = \sum_{b=1}^M \widehat{LCK}^{*b} / M.$$

6. Let $\widehat{G}_*(s) = \# \{ \widehat{LCK}^{*b} \leq s \} / M$ be the bootstrap distribution of $\{\widehat{LCK}^{*b}, b = 1, \dots, M\}$. Then a $100(1 - \alpha)\%$ bootstrap percentile interval is given by

$$(\widehat{G}_*^{-1}(\alpha/2), \widehat{G}_*^{-1}(1 - \alpha/2)).$$

A significant treatment effect is indicated if the interval does not include zero.

5. Simulation Studies

It is important to know whether the proposed bootstrap percentile interval is appropriate for practical use. Some simulation studies were performed to investigate the coverage probabilities under small sample with 10, 20, 30, and 50 per group. The tumor quadrupling times were generated from various survival distributions S_1 and S_2 of two groups with same median and given in following 3 scenarios (Su and Wei, 1993),

1. $S_1(t) = e^{-t}$ and $S_2(t) = e^{-t}$.
2. $S_1(t) = e^{-t}$ and $S_2(t) = e^{-(1.13t)1.5}$.
3. $S_1(t) = e^{-t}$ and $S_2(t) = 1 - \Phi(\log(1.44t))$, where $\Phi(\cdot)$ is a standard normal distribution function.

The censoring times are generated from uniform distribution $U(0, c_i)$ with $c_i, i = 1, 2$, which are determined by some prespecified censoring proportions at 10%, 20%, and 30%. The tumor doubling time of the control group is generated by $t_D = t_Q \times u$, where t_Q is the tumor quadrupling

time of the control and u is a uniform distribution variable of $U\left(\frac{1}{3}, \frac{2}{3}\right)$. Table 1 lists simulated empirical coverage probabilities of the bootstrap percentile interval based on 1,000 independent Monte Carlo samples and 2,000 independent bootstrap samples. The bias and standard deviation of estimated LCK from this simulation study is presented in Table 2 (results for scenario 3 is similar to the scenario 1-2 and not shown on Table 2 due to limited space). The simulation results show that the bootstrap percentile interval is slightly conservative but the coverage probabilities are reasonably close to the nominal level for practical use.

6. An Example: D456-Cisplatin Solid Tumor Xenograft Model

The original data used in this example are from a tumor xenograft experiment conducted by the Pediatric Preclinical Testing Program. In this model, the D456 human brain tumor cell line was tested for sensitivity to a single cytotoxic agent, cisplatin. Twenty mice were equally randomized to the control and treatment groups. Tumor volumes were measured weekly for a 6-week period. The raw tumor volume data are shown in Table 3. Tumor volumes quadrupled in all of the control mice on or before day 21; therefore, the mice were euthanized before the end of the study, and no tumor volumes were recorded on or after day 21. Seven mice in the treatment group had quadrupled tumor volumes on or before day 35, two mice were cured tumor by the end of study, and one mouse died of toxicity on day 21; therefore, these three mice were censored at day 42, 42 and 21, respectively. The tumor quadrupling times were calculated by using the interpolation formula given in Section 2 and listed in Table 4. The censoring times are labeled with an asterisk. The median time of tumor doubling of the control group is $\widehat{t}_d(50) = 3.96$ days. The median times of tumor quadrupling were estimated from Kaplan-Meier survival distribution functions which are $\widehat{t}_c(50) = 8.7$ days and $\widehat{t}_t(50) = 24.9$ days for control and treatment, respectively. The estimated $LCK = 1.23$. The bootstrap estimate of standard error of LCK is $\widehat{se}(LCK) = 0.22$ and the 95% bootstrap percentile interval of LCK is $(0.74, 1.65)$ which does not include zero and indicates that cisplatin has significant antitumor activity of against the D456 tumor line.

7. Conclusion and Discussion

LCK is commonly used to assess cytotoxic treatment effect in preclinical *in vivo* mouse models. Because of complicated *in vivo* tumor xenograft data, formal statistical inference of LCK has not been addressed in the literature previously. For comparisons of activity between tumors, the LCK were converted to an arbitrary activity rating. In this paper, a nonparametric bootstrap percentile interval is proposed. The antitumor activity of the cytotoxic agent can be assessed on the basis of confidence limits, and an arbitrary cuto point for LCK is therefor not needed. Furthermore, simulation results showed that the proposed bootstrap percentile interval has a good coverage probability for practical use. Finally, we point out that the tumor growth delay is not always defined in terms of tumor quadrupling times. Other thresholds can be used. For example, tumor reaches some pre-determined size. The method presented in this paper will work with other thresholds too.

Acknowledgments

Authors are thankful to the Editor and anonymous referees whose careful reading and constructive comments improved the earlier version of this article. The work was supported in part by the National Cancer Institute (NCI) support grants CA21765 and N01-CM-42216 and the American Lebanese Syrian Associated Charities (ALSAC).

References

- Collett, D. Modeling Survival data in medical research. 2nd ed. Chapman & Hall; London: 2003.
- Cox, DR.; Oakes, D. Analysis of survival data. Chapman & Hall; London: 1984.
- Corbett TH, White K, et al. Discovery and preclinical antitumor efficacy evaluations of LY32262 and LY33169. Investigational New Drugs 2003;21:33–45. [PubMed: 12795528]
- Efron B. Censored data and the bootstrap. Journal of American Statistical Association 1981;76:312–319.
- Houghton PJ, Morton CL, et al. The Pediatric Preclinical Testing Program: Description of Models and Early Testing Results. Pediatr Blood Cancer 2007;49:928–940. [PubMed: 17066459]
- Kolb EA, Gorlick R, et al. Initial Testing (Stage 1) of a Monoclonal Antibody (SCH 717454) Against the IGF-1 Receptor by the Pediatric Preclinical Testing Program. Pediatr Blood Cancer 2008;50:1190–1197. [PubMed: 18260118]

- Stuschke M, Budach V, Bamberg M, Budach W. Methods for analysis of censored tumor growth delay data. *Radiation Research* 1990;122:172–180. [PubMed: 2336463]
- Su JQ, Wei LJ. Nonparametric estimation for the difference or ratio of median failure times. *Biometrics* 1993;49:603–607. [PubMed: 8369391]
- Wu, J. Technique report. Department of Biostatistics, St. Jude Children's Research Hospital; 2007. Improved time-to-event analysis for tumor xenograft models.

Table 1

Empirical coverage probability of 95% bootstrap percentile interval of log10 cell kill

Scenario	n	Censoring proportions									
		(0,0)	(0,.1)	(0,.2)	(0,.3)	(.1,.1)	(.1,.2)	(.1,.3)	(.2,.3)		
1	10	0.973	0.965	0.961	0.948	0.969	0.962	0.947	0.949		
	20	0.969	0.968	0.968	0.965	0.972	0.962	0.963	0.971		
	30	0.970	0.968	0.963	0.956	0.973	0.975	0.973	0.968		
	50	0.954	0.965	0.963	0.959	0.957	0.958	0.958	0.952		
2	10	0.965	0.976	0.972	0.962	0.965	0.964	0.958	0.951		
	20	0.968	0.964	0.966	0.965	0.975	0.977	0.969	0.975		
	30	0.972	0.962	0.956	0.952	0.962	0.961	0.957	0.967		
	50	0.964	0.966	0.956	0.943	0.966	0.966	0.958	0.964		
3	10	0.970	0.965	0.957	0.953	0.970	0.972	0.959	0.950		
	20	0.971	0.969	0.967	0.967	0.975	0.974	0.969	0.965		
	30	0.955	0.976	0.977	0.965	0.962	0.965	0.966	0.962		
	50	0.968	0.970	0.977	0.956	0.967	0.959	0.949	0.955		

Table 2

Bias and standard deviation (std) of the estimated LCK from the simulation study

Scenario	n	LCK	Censoring proportions									
			(0,0)	(0,1)	(0,2)	(0,3)	(1,1)	(1,2)	(1,3)	(2,3)		
1	10	Bias	0.12	0.17	0.16	0.15	0.16	0.14	0.13	0.14		
		Std	0.46	0.57	0.56	0.55	0.62	0.61	0.62	0.64		
	20	Bias	0.05	0.09	0.08	0.10	0.04	0.03	0.04	0.04		
		Std	0.32	0.34	0.34	0.39	0.32	0.33	0.34	0.35		
	30	Bias	0.04	0.04	0.03	0.03	0.04	0.04	0.04	0.04		
		Std	0.25	0.25	0.26	0.28	0.27	0.28	0.28	0.29		
	50	Bias	0.02	0.02	0.02	0.01	0.02	0.03	0.03	0.03		
		Std	0.19	0.19	0.19	0.20	0.19	0.20	0.20	0.21		
2	10	Bias	0.09	0.13	0.12	0.13	0.07	0.06	0.06	0.07		
		Std	0.43	0.46	0.47	0.48	0.45	0.46	0.46	0.48		
	20	Bias	0.04	0.06	0.06	0.05	0.05	0.04	0.04	0.05		
		Std	0.25	0.28	0.29	0.29	0.28	0.28	0.29	0.29		
	30	Bias	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03		
		Std	0.21	0.21	0.21	0.21	0.23	0.24	0.24	0.25		
	50	Bias	0.01	0.02	0.01	0.02	0.01	0.01	0.01	0.01		
		Std	0.15	0.16	0.16	0.17	0.17	0.17	0.17	0.17		

Table 3

Tumor volumes (cm^3) measured in D456-cisplatin tumor xenograft model

Group	Days	Mouse									
		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Control	0	0.32	0.55	0.47	0.29	2.12	0.68	0.55	0.24	2.27	0.65
	7	0.55	2.03	1.55	0.39	6.51	2.57	2.00	0.64	8.03	2.58
	14	1.65	3.78	4.95	1.02	11.5	6.01	3.90	2.22	10.8	4.17
	21	.	.	.	2.37	.	.	.	.	.	.
Treatment	0	0.19	0.45	0.30	0.44	1.32	0.15	0.15	0.18	0.27	0.49
	7	0.15	0.70	0.58	0.36	1.46	0.08	0.22	0.15	0.81	0.79
	14	0.24	0.87	1.21	0.71	1.84	0.06	0.22	0.40	1.03	0.97
	21	0.25	1.60	1.73	1.23	2.57	0.03	0.32	0.87	2.06	1.57
	28	0.49	1.69	.	2.32	•	0.04	0.82	.	.	2.88
	35	0.05	2.43	.	.	•	0.02	.	.	.	.
	42	0.00	.	.	.	•	0.00	.	.	.	.

indicates a missing value due to mouse sacrificed after its tumor quadrupled.

• indicates that mouse died due to toxicity.

Table 4

Tumor doubling and quadrupling times (days) calculated from interpolation formula for D456-cisplatin tumor xenograft model

Group	Time	Mouse									
		M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
Control	t_D	8.0	3.7	4.1	9.9	4.3	3.7	3.8	4.9	3.8	3.5
	t_Q	12.4	7.8	8.2	15.1	10.3	7.5	8.0	9.2	9.8	7.2
Treatment	t_Q	42.0*	29.1	14.0	24.9	21.0*	42.0*	25.8	19.1	14.6	23.5

* means censoring.

t_D and t_Q are the tumor doubling and quadrupling times, respectively.