Testing Relationship between Treatment and Survival Time with an Intermediate Event[†]

Sungim Lee[1)]

Abstract

Consider a clinical trial in which the main end-point is survival. Suppose after the start of the study an intermediate event occurs which may be influenced by a covariate(or treatment). In many clinical studies the occurrence of an intermediate event may change the survival distribution. This investigation develops two-stage model which, in the first stage, models the effect of covariate on the intermediate event and models the relationship between survival time and covariate as well as the intermediate event. In this paper, the two-stage model is presented in order to model intermediate event and a test based on this model is also provided. A numerical simulations are carried out to evaluate its overall significance level.

Keywords: Survival; intermediate event; two-stage model; likelihood ratio test.

1. Introduction

An intermediate event is common and has recently got much attention in a controlled clinical study with survival as an end-point. For survival data in clinical studies refer to Lee (1994, 1998) and Lee and Sather (1995). As noticed by many researchers such as Lagakos (1976), Finkelstein and Scheonfeld (1994), Fleming *et al.* (1994), Lefkopoulou and Zelen (1995) and Nan *et al.* (2005) the intermediate event which is expected to confer survival advantage may have more important effect on survival than the treatment in many clinical situations. For example, consider the problem if chemotheraphies given to patients having cancer, have an effect on the survival. In most cases, this is also evaluated by the reduction of tumor size which may appear as the response of the treatment. Anderson *et al.* (1983) studied its relationship with survival. Another example is the analysis of Stanford heart transplant data set by Crowley and Hu (1977). They consider the time-dependent Cox regression model

$$\lambda(t \,|\, x, Y(t)) = \lambda_0(t) \exp\left\{\beta_1 x + \beta_2 Y(t)\right\}, \tag{1.1}$$

where t is the survival time, $\lambda_0(t)$ is the unspecified baseline hazard function, x is a baseline covariate and $Y(t)$ is a time-dependent indicator for the occurrence of a given intermediate event defined by

$$Y(t) = \begin{cases} 0, & t < t_Y, \\ 1, & t \geq t_Y. \end{cases}$$

[1)] Assistant professor, Department of Statistics, DanKook University, 126, Jukjeon-dong, Suji-gu, Yongin-si, Gyeonggi-do 448-701, Korea. E-mail: silee@dankook.ac.kr

Here, t_Y is the time to the event of heart transplant. They treat an intermediate event as another covariate and test its effect on the survival time. As in Crowley and Hu (1977), most of studies on intermediate event are on testing its effect to survival time by treating it as an additional covariate. And less attention is given to testing treatment effect, which surely the main goal of the entire clinical study. The same is true for series of recent work by Nam and Zelen (2001) and Lefkopoulou and Zelen (1995).

Testing treatment effect on the survival time requires much care when intermediate event exists. It is mainly because the treatment effect could affect to survival time indirectly via the intermediate event. Hence, in this paper, we propose two-stage model which, in the first stage, models the effect of covariate(or treatment) on the intermediate event and models the relationship between survival time to covariates as well as the intermediate event. As a first stage model, we simply consider the regression model, or the logistic regression model if the intermediate event is binary. For the second stage, we employ the proportional hazards model used in Crowley and Hu (1977).

We expect this two-stage model will provide a power gain in testing treatment effect. However, due to the hierarchical structure of the model, the null hypothesis we test becomes a union of simple subspaces. Further, as will be shown in later, dimension of each subspace is not same. Due to these reasons, the most beloved likelihood ratio test(LRT) has difficulty in computation and its null distribution is not understood well. In this paper, we propose an alternative test procedure which has two-stage.

The two-stage procedure(TSP) we propose combines two LRTs of each stage of the proposed two-stage model: two LRTs are obtained from the story of logistic regression model and the proportional hazards model. A main difficulty of TST arises from controlling overall significance level. To resolve this difficulty, we utilize the fact that "two LRT statistics are independent to each other".

This paper is organized as follows. In Section 2, the two-stage model is presented in order to model intermediate event and TSP is also presented. Section 3 provides the numerical simulations. The paper concludes with a discussion.

2. A Two-Stage Model and Its Procedure

2.1. A two-stage model

To assess the treatment effect on the survival time controlling for intermediate event, Nam and Zelen (2001) assume that the survival distribution for treatment belongs to a family of Lehmann's alternatives for each response level which naturally corresponds to the model in (1.1). This model treats the intermediate event as one of covariates in the model. Here, this model can be rewritten and extended to the model with interaction effect between covariate and intermediate event:

$$\lambda(t \mid x, y) = \lambda_0(t) \exp(\beta_1 x + \beta_2 y + \beta_3 xy), \tag{2.1}$$

where $y = 1$ if an immediate intermediate event occurs and otherwise, $y = 0$. Note that the time-dependent covariate $Y(t)$ in (1.1) is replaced with binary variable y since we restrict our interest on "immediate intermediate event". As mentioned in Introduction, intermediate event can occur as a response to treatment. Hence, for testing treatment
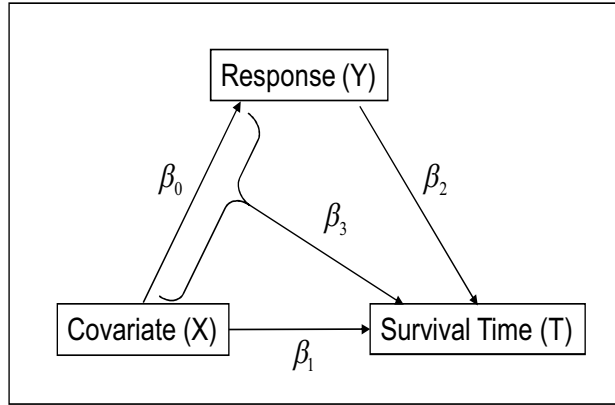
Figure 2.1: Effect path diagram: Survival model with treatment and intermediate event

effect the model will be extended as

$$f(t, y \mid x) = f(t \mid y, x)f(y \mid x), \qquad (2.2)$$

where $f(t \mid y, x)$ is modeled as in (2.1) and $f(y \mid x)$ can be assumed to have the probability function given x as

$$f(y \mid x) = \frac{\exp\left\{(\alpha + \beta_0 x)y\right\}}{1 + \exp\left(\alpha + \beta_0 x\right)}. \qquad (2.3)$$

This model may involve more interesting settings concerning the relationship between the treatment, intermediate event and survival. For example, consider the situation the survival is independent of treatment conditional on the occurrence of intermediate event. That is, $f(t \mid x, y = 1) = f(t \mid y = 1)$ and $f(t \mid x, y = 0) = f(t \mid y = 0)$ implies that the comparing treatment is done by comparing the equality of $f(y \mid x = 0)$ and $f(y \mid x = 1)$ which represents the possibility of indirect treatment effect. To test the overall effect of the treatment on the survival time, we think of a direct effect following a path $X \to T$ and an indirect effect $X \to Y \to T$ displayed in Figure 2.1. To do this in the model (2.2) applying (2.1) and (2.3), the null hypothesis can be displayed as a circle in Figure 2.2. The null hypothesis depicted as a circle($\beta_1 = \beta_3 = 0$) means no treatment effect if we only consider the model (2.1). However, this is not enough to take effect into account since a covariate could affect to the survival time indirectly via the relationship with intermediate event. Thus, the null hypothesis we are interested in should be the shaded region of the circle in Figure 2.2. One is the null situation of $H_{0A} : \beta_1 = \beta_3 = \beta_0 = \beta_2 = 0$ which implies no treatment effect on survival times conditional on the response which has no effect on the survival and this response are not affected by the treatment. The other situation is that $H_{0B} : \beta_1 = \beta_3 = \beta_0 = 0,\ \beta_2 \neq 0$ which check if the treatment affects the survival time directly and indirectly via the occurrence of intermediate event which affect the survival time. The third situation is that $H_{0C} : \beta_1 = \beta_3 = \beta_2 = 0,\ \beta_0 \neq 0$ which means that $\beta_1 = \beta_3 = 0$ still works but the treatment
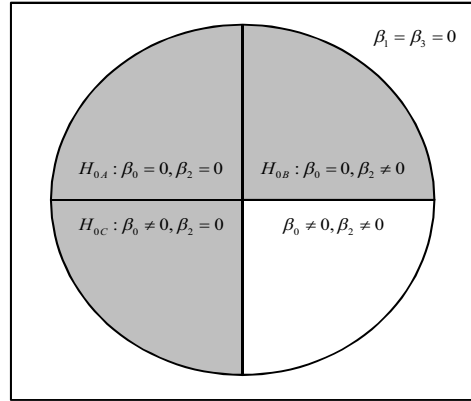
Figure 2.2: This is null parameter space which is representative of situation where there is no treatment effect on survival time

affects the occurrence of intermediate event which also affects the survival time. As a result, the interesting parameter space for testing treatment effect can be described as $\{H_{0A} \cup H_{0B} \cup H_{0C}\}$ (colored part in Figure 2.2). Here, the null hypothesis we test becomes a noncompact composite set which makes usual LRT difficult by means of commercial software. Moreover, the model in (2.1) cannot explain the indirect relationship resulting from the relationship between treatment and an intermediate event and it leads to the lose of power. Hence, we need to propose an alternative test procedures for this null parameter space: one is for testing the effect of treatment on the intermediate event and the other is for testing the effect of treatment and intermediate event on the survival time depending on the result of first procedure. In this paper we investigate these two procedures as a two-stage test procedure.

2.2. A two-stage procedure

In this section, we propose a two-stage procedure to test the treatment effect. The null hypotheses is broken down into two sequence of hypotheses and tested sequentially according to result of the first step: (i) the conditional distribution of Y given X is that in (2.3) and (ii) the conditional distribution of T given Y and X is that in (2.1). Firstly, we consider testing effect of covariate X on an intermediate event Y, *i.e.*, $H_{0A} : \beta_0 = 0$. According to the result of the first step, we will take different hypothesis. If the null hypothesis is accepted, the next step has sequential hypothesis of $H_{0B} : \beta_1 = \beta_3 = 0$. Otherwise, the sequential hypothesis will be $H_{0C} : \beta_1 = \beta_2 = \beta_3 = 0$. This step is depicted in Figure 2.3. We call this two-stage procedure in this paper. Now we introduce notations and test statistics to be used in the two-stage procedure. Let $\mathcal{H}_{0A}$, $\mathcal{H}_{0B}$ and $\mathcal{H}_{0C}$ be the set of parameter values for hypothesis, H_{0A}, H_{0B} and H_{0C} respectively. We let $\mathcal{H}_0 = \mathcal{H}_{0A} \cup \mathcal{H}_{0B} \cup \mathcal{H}_{0C}$ and let $\mathcal{H}$ be the set of all possible parameter values $(\theta, \underline{\beta})$, where $\theta = (\alpha, \beta_0)$ and $\underline{\beta} = (\beta_1, \beta_2, \beta_3)$ in (2.1) and (2.3). Let T_i and C_i be the survival time and the right-censoring time of the ith subject for $i = 1, 2, \ldots, n$ and
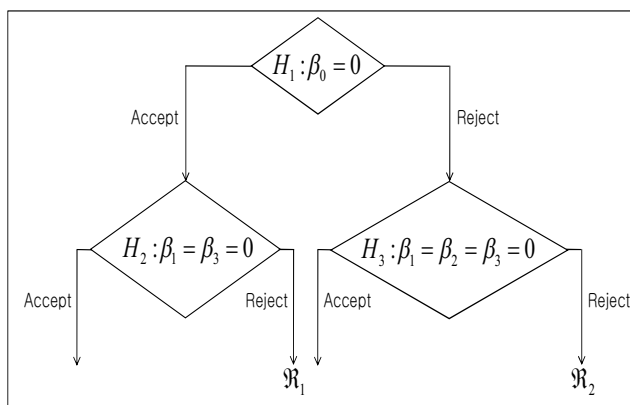
Figure 2.3: Two-stage procedure for testing the relationship between treatment and survival time

$U_i = (T_i \wedge C_i)$ and $\delta = I(T_i \wedge C_i)$. Let $\mathbf{z}_i = (x_i, y_i, x_i y_i)'$, $N_i(t) = I(U_i \leq t, \delta = 1)$ and $R_i(t) = I(T_i \geq t)$. Then, the log-likelihood function becomes

$$l\big(\alpha, \beta_0, \beta_1, \beta_2, \beta_3\big) = l\big(\theta, \underline{\beta}\big) = l_1\big(\theta\big) + l_2\big(\underline{\beta}\big), \tag{2.4}$$

where

$$l_1\big(\theta\big) = \sum_{i=1}^{n} \left\{ y_i(\alpha + \beta_0 x_i) - \log\big(1 + e^{\alpha + \beta_0 x_i}\big) \right\},$$

$$l_2\big(\underline{\beta}\big) = \sum_{i=1}^{n} \left[\int_0^\tau \mathbf{z}_i' \underline{\beta} dN_i(s) - \int_0^\tau \log\left\{ S^{(0)}(\underline{\beta}, s) \right\} dN_i(s) \right]$$

and

$$S^{(0)}\big(\underline{\beta}, t\big) = n^{-1} \sum_{j=1}^{n} R_j(t) \exp\big(\mathbf{z}_j' \underline{\beta}\big).$$

Based on the log likelihood in (2.4), we may choose the well-known LRT statistics for each hypothesis H_{0A}, H_{0B} and H_{0C}. Hence, for each hypothesis, we let the LRT as $\mathbf{V}_1$, $\mathbf{V}_2$ and $\mathbf{V}_3$ as

$$\mathbf{V}_1 = 2\left(l_1\left(\hat{\alpha}, \hat{\beta}_0\right) - l_1\left(\hat{\alpha}, 0\right) \right) \sim \chi^2(1), \tag{2.5}$$

$$\mathbf{V}_2 = 2\left(l\left(\alpha, \hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3\right) - l\left(\alpha, \hat{\beta}_0, 0, \hat{\beta}_2, 0\right) \right) \tag{2.6}$$

$$= 2\left(l_2\left(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3\right) - l_2\left(0, \hat{\beta}_2, 0\right) \right) \sim \chi^2(2),$$

$$\mathbf{V}_3 = 2\left(l\left(\alpha, \hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3\right) - l\left(\alpha, \hat{\beta}_0, 0, 0, 0\right) \right) \tag{2.7}$$

$$= 2\left(l_2\left(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3\right) - l_2\left(0, 0, 0\right) \right) \sim \chi^2(3).$$
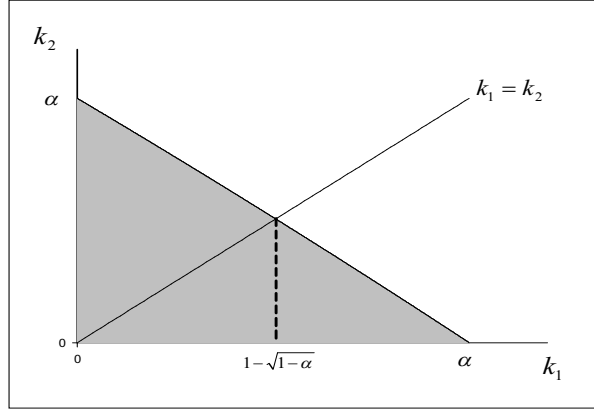
Figure 2.4: The shaded region is the feasible set $A(\alpha) = \{(k_1, k_2) : k_2 \leq \alpha - k_1/(1 - k_1)\}$ with $\gamma_1 = 1$ for the size α test

As shown in Figure 2.3, the final rejection region $\mathcal{R}_1 = \{\mathbf{V}_1 > c_1, \mathbf{V}_3 > c_3\}$ or $\mathcal{R}_2 = \{\mathbf{V}_1 < c_1, \mathbf{V}_2 > c_2\}$ is determined by the result of first stage. To do this, it however requires to control the overall significance level throughout the two-stage procedures.

Concerning the significance level in each stage, we apply the level k_1 in the first stage procedure and the level k_2 in the second stage procedure and to evaluate the significance level we generate the data from each null, H_{0A}, H_{0B} and H_{0C}. Under the data generated by the null H_{0A}, the significance level is calculated by

$$P(\mathcal{R}_1) = P(\{\mathbf{V}_1 > c_1, \mathbf{V}_3 > c_3\}) = (1 - k_1)k_2,$$
$$P(\mathcal{R}_2) = P(\{\mathbf{V}_1 < c_1, \mathbf{V}_2 > c_2\}) = k_1 k_2,$$
$$P(\mathcal{R}_1 \cup \mathcal{R}_2) = P(\mathcal{R}_1) + P(\mathcal{R}_2) = k_2. \tag{2.8}$$

Note that the test statistic $\mathbf{V}_1$ is independent of $\mathbf{V}_2$ and $\mathbf{V}_3$. We get k_2 level if the data is generated from H_{0A}. Secondly, consider the data under the null H_{0B}. In this situation, the probability of rejection region $\mathbf{V}_3 > c_3$ will be different from that in H_{0A}. Since the parameter space includes $\beta_2 \neq 0$, the probability of that region has the power γ_1. Hence, the significance level is given by

$$P(\mathcal{R}_1) = P(\{\mathbf{V}_1 > c_1, \mathbf{V}_3 > c_3\}) = (1 - k_1)k_2,$$
$$P(\mathcal{R}_2) = P(\{\mathbf{V}_1 < c_1, \mathbf{V}_2 > c_2\}) = k_1 \gamma_1,$$
$$P(\mathcal{R}_1 \cup \mathcal{R}_2) = P(\mathcal{R}_1) + P(\mathcal{R}_2) = (1 - k_1)k_2 + k_1 \gamma_1. \tag{2.9}$$

Lastly, consider the data under the null H_{0C}. As similarly in H_{0B}, the probability of rejection region $\mathbf{V}_1 > c_1$ will be the power γ_2 under $\beta_0 \neq 0$. Therefore, the significance level is calculated by

$$P(\mathcal{R}_1) = P(\{\mathbf{V}_1 > c_1, \mathbf{V}_3 > c_3\}) = (1 - \gamma_2)k_2,$$
$$P(\mathcal{R}_2) = P(\{\mathbf{V}_1 < c_1, \mathbf{V}_2 > c_2\}) = \gamma_2 k_2,$$
$$P(\mathcal{R}_1 \cup \mathcal{R}_2) = P(\mathcal{R}_1) + P(\mathcal{R}_2) = k_2. \tag{2.10}$$

Table 3.1: This numerical results show the overall significance levels across the data generation and powers at level 5% with $k_1 = k_2 = 0.025$

Sample size(n)	H_{0A} level	H_{0B} level	H_{0C} level	Power
50	0.027	0.050	0.038	0.177
100	0.031	0.064	0.034	0.536
150	0.022	0.061	0.029	0.748
200	0.023	0.052	0.041	0.861

Table 3.2: This numerical results show the overall significance levels across the data generation and powers at level 5% with $k_1 = 0$ and $k_2 = 0.05$

Sample size(n)	H_{0A} level	H_{0B} level	H_{0C} level	Power
50	0.062	0.060	0.070	0.078
100	0.063	0.053	0.052	0.069
150	0.055	0.069	0.051	0.056
200	0.046	0.047	0.063	0.060

From (2.8)–(2.10), we realize we have two kinds of significance levels according to the null. Hence, for a size α test, we can choose arbitrary significance level (k_1, k_2) satisfying

$$\max\left(k_2,\ (1 - k_1)k_2 + k_1\gamma_1\right) \le \alpha. \tag{2.11}$$

In the above, the power γ_1 will be between α and one. If we let this equal to one so that its significance level set to maximum, then the feasible set of (k_1, k_2) satisfying a size α test is shown as in Figure 2.4.

3. Numerical Simulation Studies

In this section, we numerically investigate the overall significance level and power of the proposed test at fixed level. We first introduce the set up of the simulation. We generated n survival times T which are exponentially distributed with $\lambda_0(t) = 1$ in (2.1). Censoring times are generated to yield about 30% of censored observations. Empirical significance levels and powers of the two-stage procedures were evaluated based on 1000 simulations with the data from the null H_{0A}, H_{0B} and H_{0C} respectively at the specific level k_1 and k_2. For the generation of data from the null H_{0A}, we set $\beta_0 = 0$, $\beta_2 = 0$. We also set $\beta_0 = 0$, $\beta_2 = 1$ for generation from the null H_{0B} and $\beta_0 = 1$, $\beta_2 = 0$ for the null H_{0C}. We get $\beta_0 = 1$, $\beta_2 = 1$ for evaluation of power gain. In all simulation setting, we define β_1 and β_3 equal to zero.

Case 1: If we use the same level $k_1 = k_2$ for the two-stage procedures, then the level will be $k_1 = k_2 = \sqrt{1 - \alpha}$ as in Figure 2.4. Table 3.1 is obtained from level $\alpha = 0.025$ and $k_1 = k_2 = 0.025$.

Case 2: If we consider the level only for the second stage, then this implies $k_1 = 0, k_2 = \alpha$. This implies that we only use the model (2.1). We set $\alpha = 0.05$ in Table 3.2.

Table 3.1 reports that, in most cases, the size of the test is approximately less than 0.05, the aimed significance level and there exists power gain at the indirect effect of the treatment. However, if we consider only the proportional hazards model as in (2.1), the result in Table 3.2 shows that we couldn't expect the power gain due to indirect effect although we control the overall significance level at the level 0.05.

4. Conclusion and Further Discussions

In this paper, we investigated that a covariate(say, treatment) could influence survival time indirectly via intermediate event. However, previous studies more focused on testing the intermediate event effect and do not consider the model between intermediate event and covariate which ignores the possible relationship between covariate and survival time via intermediate event. Therefore we developed the two stage model which include the model of intermediate event given a covaraite.

This paper dealt with the structures in which the waiting time for response happens relatively soon after treatment begins, provided a response is to take place. However, there are situations in which the waiting time is relative long relative to the mean survival. In this case, the longer a person is alive, the greater the change of observing a response. Although valid methods are well dealt with, the relationship between the treatment and the response has not been explored yet as we did with an immediate intermediate event. Therefore, the intermediate event having long waiting time will be explored in further studies.

References

Anderson, J. R., Cain, K. C. and Gelber, R. D. (1983). Analysis of survival by tumor response, *Journal of Clinical Oncology,* **1**, 710–719.

Crowley, J. and Hu, M. (1977). Covariance analysis of heart transplant survival data, *Journal of the American Statistical Association,* **72**, 27–36.

Finkelstein, D. M. and Scheonfeld, D. A. (1994). Analysing survival in the presence of an auxiliary variable, *Statistics in Medicine,* **13**, 1747–1754.

Fleming, T. R., Prentice, R. L., Pepe, M. S. and Glidden, D. (1994). Surrogate and auxiliary endpoints in clinical trials, with potential applications in cancer and AIDS research, *Statistics in Medicine,* **13**, 955–968.

Lagakos, S. W. (1976). A stochastic model for censored-survival data in the presence of an auxiliary variable, *Biometrics,* **32**, 551–559.

Lee, J. W. (1994). An overview of group sequential procedures, *The Korean Journal of Applied Statistics,* **7**, 35–51.

Lee, J. W. (1998). A study on the group sequential methods for comparing survival distributions in clinical trials, *The Korean Communications in Statistics,* **5**, 459–475.

Lee, J. W. and Sather, H. N. (1995). Group sequential methods for comparison of cure rates in clinical trials, *Biometrics,* **51**, 756–763.

Lefkopoulou, M. and Zelen, M. (1995). Intermediate clinical events, surrogate markers and survival, *Lifetime Data Analysis,* **1**, 73–85.

Nam, C. and Zelen, M. (2001). Comparing the survival of two groups with an interme-
diate clinical event, *Lifetime Data Analysis,* **7**, 5–19.
Nan, B., Lin, X., Lisabeth, L. D. and Harlow, S. D. (2005). A varying-coefficient Cox
model for the effect of age at a marker event on age at menopause, *Biometrics*, **61**,
576–583.