



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

Parametric Tests for Partly Interval-censored Failure Time Data under Weibull Distribution via Multiple Imputation

^{1,3}Azzah Mohammad Alharpy and ^{1,2}Noor Akma Ibrahim

¹Department of Mathematics, Faculty of Science, University Putra Malaysia,
43400 UPM Serdang, Selangor, Malaysia

²Institute for Mathematical Research, University Putra Malaysia, 43400 UPM, Malaysia

³Department of Mathematics, Faculty of Science, Taibah University, Almadinah,
Almunawwarah, Saudi Arabia

Abstract: The statistical problem considered in this article is the parametric treatment comparison when partly interval-censored failure time data exist. Partly interval-censored failure time data are composed of exact observations and interval-censored observations. This phenomenon often occurs in clinical trials and health studies that require periodic following up with patients. The authors constructed a score test and likelihood ratio test for this type of failure time data under Weibull distributions using multiple imputation technique. A simulation study and a modified secondary data set from breast cancer study are used to assess the proposed test and illustrate the differences between the two tests. The results indicate that the presented procedure works well for both tests, but the likelihood ratio test is better than the score test in certain situations.

Key words: Breast cancer study, likelihood ratio test, multiple imputation, partly interval-censored data, score test, Weibull distribution

INTRODUCTION

Partly Interval-censored (PIC) data often occurs in medical and health studies that include periodic examinations. With PIC data, the failure times are precisely observed for a portion of participants, but the failure times for remaining participants only occur within a period of time. An example of this type of data is provided by the Framingham Heart Disease study. In this study, time of first emergence of subcategory angina pectoris in coronary heart disease patients was the event of interest. For a number of patients, the event times are recorded exactly, but times are recorded only between two clinical examinations but for the remaining patients (Odell *et al.*, 1992). Research of PIC data is ongoing but limited to date. Researchers that have addressed the PIC data include Huang (1999), who derived the asymptotic properties of the nonparametric estimation for the distribution function with PIC data. Kim (2003) studied the maximum likelihood estimation for regression analysis of PIC data using the proportional hazards model. Zhao *et al.* (2008) developed a nonparametric test approach with PIC data. The nonparametric test approach is based on the same idea by Sun *et al.* (2005). Multiple Imputation (MI) has proven useful in solving many statistical problems. However, to

the best of the authors' knowledge, there is no application of MI to PIC data and the authors are investigating application to the parametric test when PIC data exist.

The objective of this article is to present parametric tests of PIC failure time data using a MI technique that can compare the survival functions of two or more samples.

PARTLY INTERVAL-CENSORED DATA

Consider schematic follow-ups for medical studies, where x_1, x_2, \dots, x_m are inspection times. Suppose that after each initial follow-up inspection times, the patient may be absent from subsequent follow-ups with probability \hat{q} , where, $0 \leq \hat{q} \leq 1$ let $T_i > 0$ be a random variable to denote failure time for i th subject. Additionally, let n be the number of participants with failure times following a continuous distribution with density function $f(t, \theta)$, where θ is a parameter vector.

Assume that the exact failure time for n_1 participants will be observed. Interval-censored failure time for the remaining subject n_2 ; $n_2 = (n - n_1)$ will be additionally observed. Exact failure times mean that any patient has the event of interest during the inspection times or the patient's condition necessitates hospital examination

where the event of interest is recorded exactly. Interval-censored failure time measurements mean that the event of interest occurs between two inspection times, (L_i, R_i) where $L_i, R_i \in (x_1, \dots, x_m)$ and $L_i < R_i$ has a probability of one. If the failure time of the patient occurs before the first examination, then left-censored will be observed, i.e., $t_i \in (0, L_i)$. If the patient did not have the failure time before or during the final examination, then right-censored will be observed, i.e., $t_i \in (R_i, \infty)$. Additionally, censoring is presumed independent of the examination time. For the interval-censored i th patient, define $\delta_i = I(t_i \in (0, L_i])$ and $\gamma_i = I(t_i \in (L_i, R_i])$. Then, the likelihood function for θ is:

$$L(\theta) = \prod_{i=1}^{n_1} f(t_i; \theta) \prod_{i=n_1+1}^n [S(L_i; \theta) - S(R_i; \theta)] \quad (1)$$

$$= \prod_{i=1}^{n_1} f(t_i; \theta) \prod_{i=n_1+1}^n [1 - S(L_i; \theta)]^{\delta_i} \left[\frac{S(L_i; \theta)}{S(R_i; \theta)} - 1 \right]^{\gamma_i} [S(R_i; \theta)]^{1-\delta_i}$$

PARAMETRIC TESTS FOR EXACT DATA

Here, briefly reviews the score test and likelihood ratio test for exact data. Suppose that P samples are generated from Weibull distribution with parameter $\theta_j = (a_j, b_j)$ and sample size n_j for j th sample, where $n = n_1 + \dots + n_p$. Assume that the survival function has the form:

$$S(t_i) = \exp \left[- \left(\frac{t_i}{b} \right)^a \right] \quad (2)$$

where, a is the shape parameter and b is the scale parameter.

The ultimate objective is to test the hypothesis $H_0: S_1(t) = \dots = S_p(t)$. The authors are interested in testing $H_0: a_1 = \dots = a_p; b_1 = \dots = b_p$ against $H_1: a_1 \neq \dots \neq a_p; b_1 \neq \dots \neq b_p$. The likelihood function under H_0 is:

$$L(a_1, b_1) = \prod_{i=1}^n \left(\frac{a_1}{b_1} \right) \left(\frac{t_i}{b_1} \right)^{a_1-1} \exp \left[- \left(\frac{t_i}{b_1} \right)^{a_1} \right] \quad (3)$$

and the likelihood function under H_1 is:

$$L(a_1, \dots, a_p, b_1, \dots, b_p) = \prod_{j=1}^p \prod_{i=1}^{n_j} \left(\frac{a_j}{b_j} \right) \left(\frac{t_{ij}}{b_j} \right)^{a_j-1} \exp \left[- \left(\frac{t_{ij}}{b_j} \right)^{a_j} \right] \quad (4)$$

To apply the score test, the score is $U(\theta)$ where:

$$U(\theta) = \frac{\partial \log L(\theta)}{\partial \theta} \quad (5)$$

The score test is based on the fact that the score has normal distribution with mean zero and variance $V(\theta)$ asymptotically under the null hypothesis, where:

$$V(\theta) = \frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta} \quad (6)$$

is the observed information matrix. The score test statistic under the null hypothesis H_0 is:

$$U^T(\hat{\theta}_0) V^{-1}(\hat{\theta}_0) U(\hat{\theta}_0) \quad (7)$$

with chi-square distribution with degree freedom (v) , where (v) is the number of imposed restrictions by the null hypothesis and $(\hat{\theta}_0) = (\hat{a}_1, \hat{b}_1)$ is the maximum likelihood estimate of θ under the null hypothesis H_0 .

For the likelihood ratio test, the test statistic is:

$$R = -2 \log R_n = -2 \log \left(\frac{\sup H_0 L_n(a_1, b_1)}{\sup H_1 L_n(a_1, \dots, a_p; b_1, \dots, b_p)} \right) \quad (8)$$

with an asymptotically χ^2 distribution with two degrees of freedom under H_0 (Kalbfleisch and Prentice, 2002).

PARAMETRIC TESTS FOR PARTLY INTERVAL-CENSORED DATA

Here, the authors display a score test and likelihood ratio test with PIC data using the MI method.

Assumption: The authors consider survival analysis which includes n independent participants from p different treatments. Let $T_i > 0$ be a random variable to signify the failure time of interest for i th participant and n_i the number of participants from treatment i with distribution function $F_i(t)$ and survival function $S_i(t) = 1 - F_i(t)$, $i = 1, \dots, p$ and $n_1 + \dots + n_p = n$. Presume that the PIC data are available for i th treatment, indicating that the exact failure time for N_i participants will be observed and given by the form:

$$\left\{ \{t_i\}_{i=1}^{N_1}; N_1 = \sum_{i=1}^p n_{1i} \right\}$$

Interval-censored failure time for the remaining participants N_2 will be observed and given by the form $\{(L_i, R_i], i = N_1+1, \dots, n; N_2 = n - N_1\}$. In this case, $(L_i, R_i]$ denotes the interval in which T_i is observed and L_i, R_i are positive random variables independent of T_i such that $L_i < R_i$ has a probability of one. When, $L_i < R_i = \infty$ the failure time T_i is regarded as a right-censored observation.

The ultimate aim is to test the hypothesis $H_0: S_1(t) = \dots = S_p(t)$ to determine whether the p treatments could have resulted from an identical failure time distribution. Let $S_0(t)$ indicate the common survival function under H_0 and let $\hat{S}_n(t)$ indicate its Parametric Maximum Likelihood Estimate (PMLE).

PMLE for partly interval-censored data: In this article, the authors suppose that T_i follows the Weibull distribution. Assume that the survival function has the form given in Eq. 2 and PIC data are available. Under these assumptions and from Eq. 1, the likelihood function is:

$$L(a, b) = \prod_{i=1}^{n_1} \left(\frac{a}{b}\right) \left(\frac{t_i}{b}\right)^{a-1} \exp\left[-\left(\frac{t_i}{b}\right)^a\right] \prod_{i=n_1+1}^n [1 - \exp\left[-\left(\frac{t_i}{b}\right)^a\right]]^{\delta_i} \left[\exp\left[\left(\frac{R_i}{b}\right)^a - \left(\frac{L_i}{b}\right)^a\right] - 1\right]^{\eta_i} \left[\exp\left[-\left(\frac{R_i}{b}\right)^a\right]\right]^{1-\delta_i} \quad (9)$$

Under the Weibull distribution, the maximum likelihood estimates of the parameters a and b are the solution of:

$$U_a(a, b) = \frac{\partial \log L(a, b)}{\partial a} = 0 \quad \text{and} \quad U_b(a, b) = \frac{\partial \log L(a, b)}{\partial b} = 0 \quad (10)$$

that can be solved by the Newton-Raphson method.

A method based on the MI technique: Here, imputation technique is used to handle the censored data. The exact failure time data will be imputed from the partial interval-censored data and applied to the parametric tests for exact data. The authors are not concerned with whether the interval-censored data are finite or infinite. The exact value can be imputed with right-censored data when the infinite interval contains part of a data set. When no data set exists inside the infinite interval, the left side of the interval will be taken as an exact value. The imputation scheme used by the authors is Rubin's MI (Rubin, 1987) and the procedure of MI is as follows. Let Y be a pre-determined positive integer and let y be an integer that satisfies $1 < y < Y$. For $y = 1, \dots, Y$ perform:

Step 1: Let T_i^y be a random sample taken from the conditional probability function:

$$P_i(t_i^y = t_k | t_i^y \in (L_i, R_i]) = \frac{\hat{S}(t_{k-1}) - \hat{S}(t_k)}{\hat{S}(L_i) - \hat{S}(R_i)}; t_k \in (L_i, R_i] \quad (11)$$

and $i = 1, \dots, N_2$; $k = 1, \dots, M$. where, $\{t\}_{k=1}^M$ denotes the unique arrangement element of $\{0, t_j, L_i, R_i$; $j = 1, \dots, N_1$, $i = N_1 + 1, \dots, n\}$

Next, a set of exact data will be obtained $\{t_i^y, i = 1, \dots, N_2\}$

Step 2: Mix the exact data imputed from the conditional probability function in step 1 with the exact data established in the original data. This step will result in exact data $\{t_i, i = 1, \dots, N_1$; $t_i^y, i = N_1 + 1, \dots, n\}$

Step 3: Apply the parametric tests, such as the score test or likelihood ratio test, for the exact data

Step 4: Repeat steps 1-3 for each $y = 1, \dots, Y$

Step 5^a: For the score test, let:

$$\bar{U} = \sum_{y=1}^Y \frac{U^y}{Y}$$

The covariance matrix V of \bar{U} is the sum of two components. The first is the average within-imputation covariance connected with U and the second is between-imputation variance of U . That is:

$$\hat{V} = \frac{1}{Y} \sum_{y=1}^Y V^y + \left(1 + \frac{1}{Y}\right) \sum_{y=1}^Y \frac{[U^y - \bar{U}][U^y - \bar{U}]^T}{Y - 1} \quad (12)$$

Hence, the suggested test statistic to compare P treatments is $\bar{U}^T (\hat{V})^{-1} \bar{U}$

The simulation result indicates that the test statistic has an approximate chi-square distribution with two degrees of freedom $\chi^2_{(2)}$ under the null hypothesis

Step 5^b: For the likelihood ratio test, let:

$$\bar{R} = \sum_{y=1}^Y \frac{R^y}{Y}$$

be the test statistic to compare P treatments. The simulation result indicates that the test statistic has an approximate chi-square distribution with two degrees of freedom $\chi^2_{(2)}$ under the null hypothesis

SIMULATION STUDY

To examine the accuracy of both tests proposed in the previous section, the authors performed a simulation study compared two treatments under a proportional hazards model. It was presumed that failure times were generated from Weibull distribution with shape parameter a and scale parameter b with survival function:

$$S_i(t) = \exp\left(-\left(\frac{t}{b}\right)^a\right)$$

for treatment 1 and $S_2(t) = (S_1(t))^{\exp(\beta)}$ for treatment 2, where β is a parameter that measures the difference between two survival functions. There are exact observations and interval-censored observations for each treatment. Additionally, the total sample size of the two treatment groups is assumed to be $n = 50, 100, 200$ and $\beta = 0.9, 0.5, 0, -0.5, -0.9$.

The following algorithm is used to generate interval-censored data:

- Generate failure time T_i ; $i = 1, \dots, n$; $n = (n_1 + n_2)$ from Weibull distribution
- Construct the set of follow-up studies that pre-specify examination times to examine participants. Each participant was assumed to be observed at 11 follow-up times which are $\tau_r = 1 + 1.5(r-1)$, $r = 1, \dots, 11$
- Assume that after the initial follow-up time, a participant may be absent from subsequent follow-ups with probability \dot{q} where $0 \leq \dot{q} \leq 1$
- Randomly choose T_i ; $i = 1, \dots, n_1$ to be exact failure time for n_1 participants such that T_i lies between the first and the last follow-up times
- The interval-censored time for the remaining n_2 participants is defined as the shortest time interval between two successful examination times which also includes the participant's generated true failure time. The left side of the interval for a participant will

be defined as zero when the true failure time occurs before the first successful examination time and the right side of the interval for a participant will be defined as infinity when the true failure time is greater than the terminal successful examination time

RESULTS AND DISCUSSION

The results obtained from simulation study are based on 1000 replications and $Y = 10$. To distinguish between two tests, let (STI) denote to the score test with MI and let (RTI) denote to the likelihood ratio test with MI. Table 1 displays the estimated power and size at the significance level of 0.05 for both tests based on the simulated data for 30% of exact data and for different sample sizes with different values of β and \dot{q} . In Table 1, values under $\beta = 0$ refer to the estimate size of test and values under $\beta = \pm 5$; ± 9 refer to the power of test. As presented in Table 1, the power of both tests decreases as \dot{q} increases, except for $\beta = 0.9$ in the score test. As anticipated, the power of both tests increases as the

Table 1: Estimated power and size of tests with 30% exact observations with 10 multiple imputations and 1000 replications at 0.05 significance level

			n = 50 (n ₁ = n ₂ = 25)			n = 100 (n ₁ = n ₂ = 50)			n = 200 (n ₁ = n ₂ = 100)		
β	a	b	q̇ = 0.1	q̇ = 0.3	q̇ = 0.5	q̇ = 0.1	q̇ = 0.3	q̇ = 0.5	q̇ = 0.1	q̇ = 0.3	q̇ = 0.5
0.9											
STI	1.5	6	0.583	0.609	0.668	0.784	0.825	0.909	0.905	0.964	0.997
LRI			0.796	0.761	0.706	0.976	0.964	0.951	1.000	1.000	1.000
STI	1	5	0.504	0.537	0.584	0.763	0.805	0.872	0.894	0.934	0.989
LRI			0.769	0.750	0.721	0.961	0.956	0.943	0.998	0.998	0.998
STI	0.6	3	0.413	0.448	0.483	0.569	0.620	0.730	0.604	0.684	0.823
LRI			0.803	0.796	0.768	0.994	0.989	0.990	1.000	1.000	1.000
0.5											
STI	1.5	6	0.428	0.389	0.329	0.604	0.577	0.505	0.878	0.858	0.788
LRI			0.318	0.304	0.260	0.577	0.554	0.506	0.867	0.855	0.811
STI	1	5	0.351	0.336	0.299	0.521	0.481	0.437	0.775	0.747	0.690
LRI			0.283	0.281	0.253	0.460	0.446	0.428	0.761	0.745	0.700
STI	0.6	3	0.338	0.323	0.293	0.564	0.555	0.515	0.814	0.800	0.778
LRI			0.296	0.271	0.257	0.539	0.531	0.505	0.812	0.806	0.796
0											
STI	1.5	6	0.082	0.069	0.062	0.061	0.051	0.041	0.042	0.033	0.029
LRI			0.045	0.047	0.041	0.049	0.043	0.039	0.035	0.037	0.038
STI	1.0	5	0.076	0.072	0.069	0.055	0.051	0.038	0.035	0.028	0.024
LRI			0.046	0.048	0.048	0.041	0.037	0.037	0.031	0.028	0.028
STI	0.6	3	0.088	0.092	0.072	0.057	0.051	0.050	0.033	0.030	0.022
LRI			0.050	0.047	0.051	0.042	0.042	0.043	0.028	0.026	0.021
-0.5											
STI	1.5	6	0.247	0.222	0.183	0.374	0.342	0.295	0.632	0.597	0.519
LRI			0.182	0.181	0.104	0.327	0.321	0.295	0.613	0.602	0.550
STI	1.0	5	0.153	0.154	0.128	0.228	0.215	0.186	0.420	0.391	0.344
LRI			0.117	0.118	0.104	0.195	0.194	0.174	0.405	0.381	0.362
STI	0.6	3	0.140	0.131	0.114	0.162	0.156	0.124	0.250	0.227	0.193
LRI			0.098	0.091	0.082	0.135	0.129	0.114	0.241	0.214	0.200
-0.9											
STI	1.5	6	0.438	0.432	0.393	0.764	0.730	0.658	0.964	0.958	0.928
LRI			0.402	0.389	0.348	0.722	0.711	0.657	0.959	0.957	0.941
STI	1.0	5	0.290	0.275	0.255	0.514	0.496	0.446	0.823	0.806	0.759
LRI			0.236	0.221	0.204	0.466	0.474	0.429	0.812	0.798	0.773
STI	0.6	3	0.294	0.285	0.272	0.435	0.421	0.373	0.683	0.654	0.599
LRI			0.216	0.216	0.216	0.382	0.374	0.345	0.663	0.638	0.601

sample size increases, indicating different effects of n and \hat{q} on the power of both tests. The \hat{q} has a smaller effect than the sample size on the power of both tests. Table 1 additionally confirms that the power of both tests increases as the absolute value of β increases. According to whether the estimated size of test is close to or far from the significant level of 0.05, Table 1 showed that the estimated size of the test in the RTI is more accurate than the STI when $n = 50$. Meanwhile, the estimated size of the test in the STI is more accurate than the RTI when $n > 50$. The table reveals that both tests work well under this situation considered. Nevertheless, RTI is better than STI in certain cases. The authors studied the estimated power and size at the significance level of 0.05 for both tests based on the simulated data for 50% of exact data and different sample sizes with different values of β and \hat{q} . The results are similar to Table 1. The estimated size of both tests improved in the majority of cases when more exact data exists.

AN APPLICATION

The authors conducted several modifications on the breast cancer study presented by Klien and Moeschberger (1997) to test the suitability of the new tests. Several exact observations were added by generating random failure times that lie within the same range of the given data. The data set are shown in Table 2.

The data set consists of 124 participants, of whom 61 patients were treated by radiation therapy (R) and 63 patients were treated by radiation therapy and adjuvant chemotherapy (R+C). This study was executed to compare the cosmetic effects of R against R+C on women with early breast cancer and the event of interest was the time to the first occurrence of breast retraction. The patients were observed at clinic visits every 4 or 6 months and the actual dates of the event were recorded exactly if available. The interval of events was noted when actual dates were unavailable. The target of this study is to compare two treatments with respect to their cosmetic effects.

In both tests, STI and RTI were applied to examine the cosmetic effect between two treatments. The obtained values of the test statistic are equal to 7.655 and 7.262, respectively, with p-values of 0.022 and 0.026, respectively. These results indicate a significant cosmetic effect difference between the treatments. In addition, the STI p-values are smaller than the RTI p-values, indicating that the STI is more sensitive in detecting the difference between the treatments. Notably, the sample size is large where $n = 124$. Therefore, it can be easily observed that

Table 2: Time to cosmetic deterioration (in months) in breast cancer patients who received two treatments

(0, 7]; (0, 8]; (0, 5]; (4, 11]; (5, 12]; (5, 11]; (6, 1 0]; (7, 16]; (7, 14]; (11, 15]; (11, 18]; ≥ 15 ; ≥ 17 ; (17, 25]; (17, 25]; ≥ 18 ; (19, 35]; (18, 26]; ≥ 22 ; ≥ 24 ; ≥ 24 ; (25, 37]; (26, 40]; (27, 34); ≥ 32 ; ≥ 33 ; ≥ 34 ; (36, 44]; (36, 48]; ≥ 36 ; (37, 44]; ≥ 37 ; ≥ 37 ; ≥ 37 ; ≥ 38 ≥ 40 ; ≥ 45 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 46 ; ≥ 36 ; 2; 15; 17; 4; 18; 20; 21; 28; 30; 34; 22; 40; 42; 46; 36	(0, 22]; (0, 5]; (4, 9]; (4, 8]; (5, 8]; (8, 12]; (8, 21]; (10, 35] (10, 17]; (11, 13]; ≥ 11 ; (11, 17]; ≥ 11 ; (11, 20]; (12, 20]; ≥ 13 ; (13, 39]; ≥ 13 ; ≥ 13 ; (14, 17]; (14, 19]; (15, 22]; (16, 24]; (16, 20]; (16, 24]; (16, 60]; (17, 27]; (17, 23] (18, 25]; (18, 24]; (19, 32]; ≥ 21 ; (22, 32]; (17, 26]; ≥ 23 ; (24, 31]; (24, 30]; (32, 40]; (30, 36]; ≥ 31 ; ≥ 32 ; (32, 40); ≥ 34 ; ≥ 34 ; ≥ 35 ; (35, 39]; (44, 48]; ≥ 48 ; 4; 7; 10; 13; 15; 19; 22; 14; 18; 25; 29; 30; 32; 34; 40
---	--

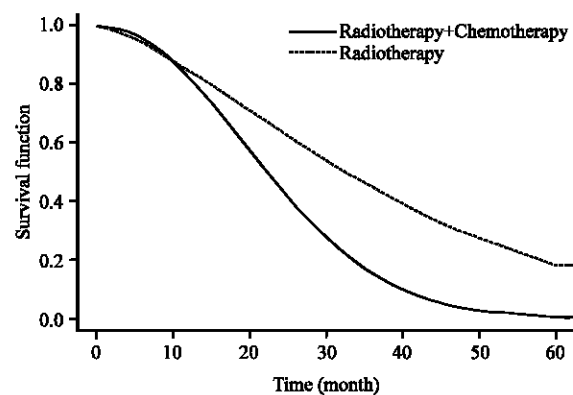


Fig. 1: PMLE of time survival function to cosmetic deterioration

this result coincide with the simulation results. Furthermore, Fig. 1 shows that the Weibull distribution is used to obtain estimated survival function for the two treatments R and R+C. This figure indicates that patients in the R+C group develop breast retraction earlier than patients in the R group.

CONCLUSION

In this article, the authors proposed parametric tests using the MI technique, such as STI and RTI for PIC failure time data under Weibull distribution to compare survival functions of p treatments. Simulation results indicate that the tests work well under the situations considered in the study. The test comparison revealed that the power of the test is affected by the accuracy of the estimated test size in both tests. Therefore, the performance of a test, effective or otherwise, can be evaluated according to the estimated size of the test. Hence, it was found that RTI is better than STI when the sample size is equal to 50 and STI is better than RTI when the sample size is greater than 50. Additionally, the proposed parametric tests have been used successfully to detect the difference between the R and R+C treatments.

REFERENCES

- Huang, J., 1999. Asymptotic properties of nonparametric estimation based on partly interval-censored data. *Stat. Sin.*, 9: 501-520.
- Kalbfleisch, J.D. and R.C. Prentice, 2002. *The Statistical Analysis of Failure Time Data*. 2nd Edn., John Wiley and Sons, New Jersey.
- Kim, J.S., 2003. Maximum likelihood estimation for the proportional hazards model with partly interval-censored data. *J. R. Stat. Soc.: Ser. B Stat. Methodol.*, 65: 489-502.
- Klien, J.P. and M.L. Moeschberger, 1997. *Survival Analysis Techniques for Censored and Truncated Data*. 2nd Edn., Springer-Verlag, New York.
- Odell, P.M., K.M. Anderson and R.B. D'Agostino, 1992. Maximum likelihood estimation for interval-censored data using a Weibull-based accelerated failure time model. *Biometrics*, 48: 951-959.
- Rubin, D.B., 1987. *Multiple Imputation for Nonresponse in Surveys*. John Wiley and Sons, New York.
- Sun, J., Q. Zhao and X. Zhao, 2005. Generalized log-rank tests for interval-censored failure time data. *Scand. J. Stat.*, 32: 49-57.
- Zhao, X., Q. Zhao, J. Sun and J.S. Kim, 2008. Generalized log-rank tests for partly interval-censored failure time data. *Biom. J.*, 50: 375-385.