Biostatistics (2013), **14**, 3, *pp.* 422–432
doi:10.1093/biostatistics/kxt002
Advance Access publication on February 19, 2013

Estimating incremental cost-effectiveness ratios and their confidence intervals with different terminating events for survival time and costs

SHUAI CHEN*

Department of Statistics, Texas A&M University, College Station, TX 77843, USA
shuai@stat.tamu.edu

HONGWEI ZHAO

Department of Epidemiology and Biostatistics, Texas A&M Health Science Center, College Station, TX 77843, USA

Summary

Cost-effectiveness analysis (CEA) is an important component of the economic evaluation of new treatment options. In many clinical and observational studies of costs, censored data pose challenges to the CEA. We consider a special situation where the terminating events for the survival time and costs are different. Traditional methods for statistical inference offer no means for dealing with censored data in these circumstances. To address this gap, we propose a new method for deriving the confidence interval for the incremental cost-effectiveness ratio. The simulation studies and real data example show that our method performs very well for some practical settings, revealing a great potential for application to actual settings in which terminating events for the survival time and costs differ.

Keywords: Censored data; Cost-effectiveness analysis; Different terminating events; Fieller method; Survival analysis.

1. Introduction

With health-care costs surging in an environment of limited resources, economic evaluations of new treatment strategies are becoming more and more prevalent. If a program has higher cost but greater benefit than its competitor, a decision must be made on which of the two programs to adopt. The incremental cost-effectiveness ratio (ICER) is designed to measure the trade-off between the costs and health benefits of medical interventions. It is defined as the extra costs incurred for saving one additional year-of-life (YOL). The ICER has been the most widely used tool for cost-effectiveness analysis (CEA) (Zwanziger *and others*, 2006; Wailoo *and others*, 2008; McIntosh *and others*, 2009).

Analyzing cost data requires advanced statistical methodologies, especially when cost data are censored. Over a decade has passed since scholars recognized that caution should be exercised when using censored cost data, as described in Lin *and others* (1997). The authors point out that traditional methods for handling

*To whom correspondence should be addressed.

censored survival data, such as the Kaplan–Meier estimator, Log-rank test, and Cox proportional hazards regression are no longer valid for analyzing censored cost data, due to the "induced informative censoring" problem. Additionally, because of censoring, the costs and survival distribution cannot be estimated over the entire health history, unless more assumptions are imposed. Hence, a limited time horizon, such as L (years), is often required, i.e. we measure life-years saved within a limited horizon L, costs within L, and hence ICER within L.

Since the ICER is a ratio statistic with quite a skewed distribution, authors often construct a confidence interval (CI) for the ICER in order to estimate its variability. Researchers have proposed various methods for how to find CIs for the ICER. The most widely used in health service research and the health economic literature are bootstrap methods (Efron and Tibshirani, 1986, 1993; Hwang, 1995; Mushlin *and others*, 1998; Jiang *and others*, 2000; O'Brien and Briggs, 2002; Jiang and Zhou, 2004); but one can also adapt Fieller's theorem to censored cost data to obtain the CI for the ICER (Fieller, 1954; Chaudhary and Sterns, 1996; Zhao and Tian, 2001; Wang and Zhao, 2008). Although many researchers believe that since the Fieller method is based on the large sample normal assumption, the bootstrap methods provide better coverage, Hwang (1995) and Jiang *and others* (2000) showed that both methods are first-order accurate. Therefore, the Fieller method, if used correctly, can be a reliable and efficient way to compute these CIs.

To obtain the CI for the ICER using Fieller's theorem, we need to estimate not only the mean costs and effectiveness (e.g. life expectancies) and their respective variances, but also their covariance. There are proposed methods for estimating the mean medical costs and related variance, and most of these focus on the time-restricted medical costs (Lin *and others*, 1997; Bang and Tsiatis, 2000; Zhao and Tian, 2001; O'Hagan and Stevens, 2004; Raikou and McGuire, 2004; Zhao *and others*, 2007, among others). In the construction of the CI, a challenge caused by an earlier stopping time for cost collection for some patients has been addressed by Wang and Zhao (2006). However, another challenge arises when the terminating events for costs and survival are different. For example, in the Multicenter Automatic Defibrillator Implantation Trial with Cardiac Resynchronization Therapy (MADIT-CRT), the primary goal was to determine whether the CRT with biventricular pacing would reduce the risk of death or heart failure (HF) events in patients with mild cardiac symptoms (Moss *and others*, 2009). The terminating event for the effectiveness measure is death or HF, whichever occurs first. Meanwhile, some patients who experienced HF events were still living, and continued to accumulate costs and report these costs. Since the treatment costs accumulated until death is of interest, the terminating event for this cost evaluation is a different one, death. We anticipate that more and more medical advances will occur, which may not prolong the overall survival time, but will prevent adverse events such as HFs. Therefore, ICERs with different terminating events might be used more frequently in the future.

The remainder of the article is organized as follows. In Section 2, we propose a method for estimating the ICER and constructing its CI with censored data and different terminating events. In Section 3, we perform numerical studies to examine the empirical coverage probability of the CI for this special ICER, and compare its performance with bootstrap methods. Next, we illustrate our method by applying it to the MADIT-CRT study. Finally, we provide discussions and concluding remarks.

2. Methods

2.1 *Notation and assumptions*

We first concentrate on patients in one arm of the study. For the ith person, let T_i denote the overall survival time, i.e. time until death. In addition, the subject may experience a HF event at time HF_i. Let T_i^{F} denote the HF-free survival time, i.e. the time to a HF event or death, whichever occurs first, $T_i^{\mathrm{F}} = \min(\mathrm{HF}_i, T_i)$. Let C_i represent the censoring time. Define the observed follow-up time as $X_i = \min(T_i, C_i)$, and the death indicator as $\Delta_i = I(T_i \leqslant C_i)$, where $I(\cdot)$ is the indicator function. Similarly, define the HF-free follow-up

time as $X_i^{\mathrm{F}} = \min(T_i^{\mathrm{F}}, C_i)$, and the HF-free survival event indicator as $\Delta_i^{\mathrm{F}} = I(T_i^{\mathrm{F}} \leqslant C_i)$. Let $M_i(u)$ be the costs accumulated over time u. For simplicity, we define $M_i = M_i(X_i)$ as the observed total costs.

We assume that the censoring time C_i is independent of the survival time T_i, the HF time HF_i, and the cost history process $\{M_i(u), u \leqslant T_i\}$. This assumption is reasonable for a well-conducted clinical trial. Because of censoring, it is impossible to estimate the costs over the entire health history. Therefore, we consider only costs accumulated up to a maximum of L units of time, where L is chosen based on the availability of data. This is equivalent to redefining our survival time as $T_i^* = \min(T_i, L)$, and $T_i^{\mathrm{F}*} = \min(T_i^{\mathrm{F}}, L)$. For ease of notation, we suppress the superscript $*$ of T_i^* and $T_i^{\mathrm{F}*}$ throughout the paper.

For each of the two treatment groups, $k = 0, 1$, we observe the following identically distributed, independent data $\{X_i, \Delta_i, X_i^{\mathrm{F}}, \Delta_i^{\mathrm{F}}, M_i(X_i),\ i = 1, \ldots, n_k\}$; n_k is the number of patients for arm k. Our goal is to estimate the mean cost $\mu^{\mathrm{M}} = E\{M_i(T_i)\}$ and the mean HF-free survival time $\mu^{\mathrm{F}} = E(T_i^{\mathrm{F}})$ for each of the treatment groups, and to compare the treatment strategies by obtaining the ICER and its CI, based on the estimated differences of μ^{M} and μ^{F} from the two groups and their variances and covariances.

2.2 *Estimating the mean costs for each group*

Here, we briefly review the methods for estimating the mean costs accumulated over time L with censored data. Bang and Tsiatis (2000) proposed a consistent estimator based on the inverse probability weighting technique: $\hat{\mu}_{\mathrm{BT}}^{\mathrm{M}} = (1/n)\sum_{i=1}^{n}(\Delta_i M_i / \hat{K}(T_i))$, where $\hat{K}(T_i)$ is the Kaplan–Meier estimator for the survival function of the censoring time C, $K(u) = \Pr(C_i > u)$. This is the simple unpartitioned version, and Bang and Tsiatis (2000) also provided a partitioned estimator BTp.

When the cost history is available, the BT estimator is not efficient since it does not use the cost information from censored observations. A more efficient estimator, which is also easy to use, is proposed by Zhao and Tian (2001). The ZT estimator has the following simplified form (Pfeifer and Bang, 2005):

$$\hat{\mu}_{\mathrm{ZT}}^{\mathrm{M}} = \frac{1}{n}\sum_{i=1}^{n}\frac{\Delta_i M_i}{\hat{K}(T_i)} + \frac{1}{n}\sum_{i=1}^{n}\frac{(1-\Delta_i)\{M_i(C_i) - \overline{M(C_i)}\}}{\hat{K}(C_i)}, \tag{2.1}$$

where $\overline{M(C_i)} = \sum_{j=1}^{n} I(X_j \geqslant C_i) M_j(C_i) / \sum_{j=1}^{n} I(X_j \geqslant C_i)$, which is the average accumulated costs at time C_i of those subjects who are alive at C_i.

Zhao *and others* (2007) described the conditions under which the ZT estimator is equivalent to the BTp estimator, as well as the two estimators LinA/B proposed by Lin *and others* (1997).

Zhao and Tian (2001) show that the ZT estimator is consistent, and asymptotically normally distributed with variance that can be estimated consistently by

$$\begin{aligned}\hat{\mathrm{Var}}(\hat{\mu}_{\mathrm{ZT}}^{\mathrm{M}}) = {} & \frac{1}{n^2}\sum_{i=1}^{n}\frac{\Delta_i(M_i - \hat{\mu}_{\mathrm{ZT}}^{\mathrm{M}})^2}{\hat{K}(T_i)} + \frac{1}{n^2}\int_0^L \frac{\mathrm{d}N^C(u)}{\hat{K}(u)^2}\{\hat{G}(M^2, u) - \hat{G}(M, u)^2\} \\ & - \frac{2}{n^2}\int_0^L \frac{\mathrm{d}N^C(u)}{\hat{K}(u)^2}[\hat{G}\{MM(u), u\} - \hat{G}(M, u)\hat{G}\{M(u), u\}] \\ & + \frac{1}{n^2}\int_0^L \frac{\mathrm{d}N^C(u)}{\hat{K}(u)^2}[\hat{G}^*\{M(u)^2, u\} - \hat{G}^*\{M(u), u\}^2],\end{aligned} \tag{2.2}$$

where $N^C(u) = \sum_{i=1}^{n} N_i^C(u) = \sum_{i=1}^{n} I(X_i \leqslant u, \Delta_i = 0)$, $\hat{G}^*\{Z, u\} = \left\{\sum_{i=1}^{n} Z_i Y_i(u)\right\} / Y(u)$, $Y(u) = \sum_{i=1}^{n} Y_i(u) = \sum_{i=1}^{n} I(X_i \geqslant u)$, and $\hat{G}(Z, u) = (1/n\hat{S}(u))\sum_{i=1}^{n}(\Delta_i/\hat{K}(T_i))Z_i I(T_i \geqslant u)$ for any random variable Z, and $\hat{S}(u)$ is the Kaplan–Meier estimator for $S(u)$, the survival distribution of T at time u, using

data $(X_i, \Delta_i,\ i = 1, \ldots, n)$. This variance formula is a simplified form of the original formula given by Zhao and Tian (2001) and Zhao and Wang (2010).

2.3 *Estimating the mean HF-free survival time for each group*

The mean survival time up to time L can be obtained by the area under the survival function, i.e. $\hat{\mu}^T = \int_0^L \hat{S}(x)dx$, where $\hat{S}(x)$ is the Kaplan–Meier estimator for $S(u) = \Pr(T > u)$. This estimator can be more conveniently obtained (Satten and Datta, 2001; Zhao and Tian, 2001) by

$$\hat{\mu}^T = \frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i T_i}{\hat{K}(T_i)}. \tag{2.3}$$

Similarly, the mean HF-free survival time can be estimated by

$$\hat{\mu}^{\mathrm{F}} = \frac{1}{n}\sum_{i=1}^{n} \frac{\Delta_i^{\mathrm{F}} T_i^{\mathrm{F}}}{\hat{K}^{\mathrm{F}}(T_i^{\mathrm{F}})}, \tag{2.4}$$

where $\hat{K}^{\mathrm{F}}(u)$ is the Kaplan–Meier estimator for $K(u) = \Pr(C > u)$, the survival distribution of C at time u, using data $(X_i^{\mathrm{F}}, \Delta_i^{\mathrm{F}},\ i = 1, \ldots, n)$. Following Zhao and Tian (2001), its variance can be estimated consistently by

$$\frac{1}{n^2}\sum_{i=1}^{n} \frac{\Delta_i^{\mathrm{F}}(T_i^{\mathrm{F}} - \hat{\mu}^{\mathrm{F}})^2}{\hat{K}^{\mathrm{F}}(T_i)} + \frac{1}{n^2}\int_0^L \frac{\mathrm{d}N^{\mathrm{F}}(u)}{\hat{K}^{\mathrm{F}}(u)^2}[\hat{G}^{\mathrm{F}}\{(T^{\mathrm{F}})^2, u\} - \hat{G}^{\mathrm{F}}(T^{\mathrm{F}}, u)^2],$$

where $N^{\mathrm{F}}(u) = \sum_{i=1}^n N_i^{\mathrm{F}}(u) = \sum_{i=1}^n I(X_i^{\mathrm{F}} \leqslant u, \Delta_i^{\mathrm{F}} = 0)$, $\hat{G}^{\mathrm{F}}(Z, u) = (1/n\hat{S}^{\mathrm{F}}(u))\sum_{i=1}^n(\Delta_i^{\mathrm{F}}/\hat{K}^{\mathrm{F}}(T_i^{\mathrm{F}}))$ $Z_i I(T_i^{\mathrm{F}} \geqslant u)$, $\hat{S}^{\mathrm{F}}(u)$ is the Kaplan–Meier estimator for $S^{\mathrm{F}}(u) = \Pr(T_i^{\mathrm{F}} > u)$.

2.4 *Estimating the ICER and its CI*

The ICER is the ratio of the difference of costs and the difference of effects between two treatment groups. Here, we use HF-free survival time as the measure of effectiveness. For a two-arm trial, for each group k $(k = 0, 1)$, denote μ_k^{M} as the mean cost and μ_k^{F} as the mean HF-free survival time, each limited to a window of time $[0, L]$. The ICER, which measures the additional costs needed for saving 1 year of HF-free lifetime, is defined as $\gamma = (\mu_1^{\mathrm{M}} - \mu_0^{\mathrm{M}})/(\mu_1^{\mathrm{F}} - \mu_0^{\mathrm{F}})$.

The ICER γ can be estimated by plugging in the ZT estimator (2.1) for the mean cost $\hat{\mu}_k^{\mathrm{M}}$, and the estimator (2.4) for the mean HF-free survival time, $\hat{\mu}_k^{\mathrm{F}}$, for each group $k, k = 0, 1$. We use Fieller's Theorem to obtain CIs for the ICER, similarly to Zhao and Tian (2001). Since asymptotically $x = \hat{\mu}_1^{\mathrm{M}} - \hat{\mu}_0^{\mathrm{M}}$ and $y = \hat{\mu}_1^{\mathrm{F}} - \hat{\mu}_0^{\mathrm{F}}$ are bivariate normally distributed, the $100(1 - 2\alpha)\%$ confidence limits for the ICER γ are

$$\frac{xy - z_\alpha^2 s_{xy} \pm \{(xy - z_\alpha^2 s_{xy})^2 - (x^2 - z_\alpha^2 s_{xx})(y^2 - z_\alpha^2 s_{yy})\}^{1/2}}{y^2 - z_\alpha^2 s_{yy}}, \tag{2.5}$$

where s_{xx} and s_{yy} are respectively the variances of x and y, s_{xy} is the covariance of x and y, and z_α is the cut-off point with tail area α of the standard normal distribution. If the denominator of (2.5) is positive, the CI is finite. If the denominator of (2.5) is negative, meaning that the difference between the effects of two treatments is not statistically significant, the CI for the ICER is exclusive and thus infinite. For a discussion on the interpretation of infinite intervals, see Wang and Zhao (2008).

The variances of x and y, s_{xx} and s_{yy}, can be obtained from the results mentioned above, treating two arms as independent samples. The challenge is to find the covariance between x and y, or the covariance between the mean cost estimator and the mean HF-free survival time estimator $\hat{\mu}_k^{\mathrm{M}}$ and $\hat{\mu}_k^{\mathrm{F}}$, due to the different terminating events used here. In the Appendix of supplementary material available at *Biostatistics* online, we express the mean HF-free survival time estimator $\hat{\mu}_k^{\mathrm{F}}$ and the mean cost estimator $\hat{\mu}_k^{\mathrm{M}}$ in martingale forms, and derive the covariance between them based on the counting process theory and the general theory for the missing data process (Fleming and Harrington, 1991; Robins and Rotnitzky, 1992; Robins *and others*, 1994). We show that the covariance between $\hat{\mu}_k^{\mathrm{M}}$ and $\hat{\mu}_k^{\mathrm{F}}$ can be estimated consistently by

$$\begin{aligned}\hat{\mathrm{Cov}}(\hat{\mu}^{\mathrm{M}}, \hat{\mu}^{\mathrm{F}}) = &\frac{1}{n^2}\sum_{i=1}^{n}\frac{\Delta_i M_i T_i^{\mathrm{F}}}{\hat{K}(T_i)} - \frac{1}{n^3}\sum_{i=1}^{n}\frac{\Delta_i M_i}{\hat{K}(T_i)}\sum_{i=1}^{n}\frac{\Delta_i^{\mathrm{F}} T_i^{\mathrm{F}}}{\hat{K}^{\mathrm{F}}(T_i^{\mathrm{F}})}\\ &+ \frac{1}{n^2}\int_0^L \frac{\mathrm{d}N^{\mathrm{F}}(u)}{\hat{K}^{\mathrm{F}}(u)^2}\{\hat{G}^{\mathrm{F}_0}(T^{\mathrm{F}}M, u) - \hat{G}^{\mathrm{F}_0}(M, u)\hat{G}^{\mathrm{F}_0}(T^{\mathrm{F}}, u)\}\\ &- \frac{1}{n^2}\int_0^L \frac{\mathrm{d}N^{\mathrm{F}}(u)}{\hat{K}^{\mathrm{F}}(u)^2}[\hat{G}^{\mathrm{F}_0}\{T^{\mathrm{F}}M(u), u\} - \hat{G}^{\mathrm{F}_0}\{M(u), u\}\hat{G}^{\mathrm{F}_0}(T^{\mathrm{F}}, u)], \qquad (2.6)\end{aligned}$$

where $\hat{G}^{\mathrm{F}_0}(Z, u) = (1/n\hat{S}^{\mathrm{F}}(u))\sum_{i=1}^{n}(\Delta_i/\hat{K}(T_i))Z_i I(T_i^{\mathrm{F}} \geqslant u)$.

3. SIMULATION

We conduct simulation studies to examine the performance of the covariance formula, and the coverage probability of the CI of the ICER that relies on the covariance formula. The overall survival time has an exponential distribution $T \sim \exp(10)$, which is the same for both treatment groups. The HF time is also exponentially distributed, but is different for each group: HF $\sim \exp(6)$ for Group 0, and HF $\sim \exp(12)$ for Group 1. Hence, Group 1 represents a new treatment which prevents the occurrence of HFs but is not effective in preventing death. The survival time T and HF time HF are generated independently and truncated at L=10. The HF-free survival time is defined as $T^{\mathrm{F}} = \min(T, \mathrm{HF})$, as mentioned previously. The censoring time has a uniform distribution, $C \sim \mathrm{Unif}(0, 15)$, resulting in 42% censoring for the overall survival time and 24–30% censoring for the HF-free survival time. The true mean HF-free survival time is 3.49 and 4.58 for Group 0 and Group 1, respectively.

We consider U-shaped sample paths for the costs, similarly to Bang and Tsiatis (2000) and Zhao *and others* (2012). The entire time period [0, 10] is partitioned into 10 equal (yearly) intervals. The total costs consist of the initial costs incurred at the beginning of the study, the terminal costs accumulated during the last year before death, the fixed annual costs which do not change for each patient, and the random annual costs which vary from year to year. We consider two scenarios with uniformly distributed costs and log-normally distributed costs. For the uniform setting, the initial costs, fixed annual costs, random annual costs, and terminal costs are uniformly distributed in (1,000, 3,000), (2,000, 4,000), (0,400), (5,000, 15,000) for Group 0, and in (20,000, 30,000), (2,000, 3,000), (0,400), (5,000, 15,000) for Group 1. For the log-normal setting, these costs are log-normally distributed with parameters (8.5, 0.632^2), (8, 0.245^2), (4, 0.245^2), and (8, 0.632^2) for Group 0, and (10.2, 0.632^2), (7.5, 0.245^2), (4, 0.245^2), and (8, 0.632^2) for Group 1. The true mean costs for Group 0 and Group 1 are \$28 241 and \$48 095 under the uniform setting, and \$27 932 and \$47 133 under the log-normal setting. Thus, Group 1 is associated with a longer HF-free survival time, but is also more costly than Group 0, mainly due to large initial costs.

The simulation results based on 2000 runs, and various sample sizes, are summarized in Table 1. We first examine the performance of the covariance estimator. Here SCov represents the sample covariance of

Table 1. *Summary of covariance estimation for costs and HF-free survival time, and the empirical coverage probability of CI for the ICER for different nominal levels* (0.95, 0.90, 0.80) *from* 2000 *simulations*

			Covariance		Empirical coverage probability for ICER		
Cost	Sample size	Group	SCov	ECov	0.95	0.90	0.80
Uniform	100	0	187	176	0.938	0.887	0.784
		1	182	193			
	200	0	82	83	0.948	0.897	0.789
		1	107	96			
	400	0	39	41	0.942	0.898	0.794
		1	46	48			
Log normal	100	0	177	196	0.934	0.889	0.788
		1	179	164			
	200	0	95	94	0.946	0.894	0.786
		1	75	81			
	400	0	47	46	0.953	0.897	0.790
		1	39	40			

SCov is the sample covariance of the mean cost estimator and the mean HF-free survival estimator; ECov is the mean of estimated covariance.

mean costs and mean HF-free survival times, and ECov represents the mean of the estimated covariance using our formula (2.6). As expected, the estimated covariances between the cost estimator and the HF-free survival estimator are very close to the sample covariances.

For each run of the simulation study, we calculate also the ICER, the extra costs incurred for saving a year of HF-free survival time, and its CI, within a time limit of L=10 years. We then calculate the coverage probability of the true ICER by the CI over the 2000 simulations, which is also shown in Table 1. We see that, for various nominal levels, the empirical coverage probability is very close to the nominal level, especially when the sample size is large.

At the suggestion of one reviewer, we also examined the performance of the bootstrap CIs. For our simulation scenarios, the bootstrap samples lie in the Northeast and Northwest regions of the cost-effectiveness (CE) plane. We consider a naive way of constructing the $100(1-2\alpha)\%$ CI using the bootstrap percentile method, i.e. arranging the bootstrap ICERs in ascending order, and obtain the $100(1-2\alpha)\%$ CI using the upper and lower $100\alpha\%$ cut-off points. We also examine the reordered bootstrap method proposed by Wang and Zhao (2008), where the orders of ICER are rearranged according to their positions from a CE plane before obtaining the tail cut-off points. The comparison of the three methods is shown in Table 2. The bootstrap percentile has much higher coverage probabilities than the nominal levels, and this poor performance is due to the fact that it always produces finite intervals, which is incorrect when the difference of the effectiveness between two groups is non-significant. The performance of the reordered bootstrap method is comparable with our method. However, our method runs much faster than the two bootstrap methods (1 min vs 8 h).

4. A real data example: MADIT-CRT

In MADIT-CRT study, patients were recruited into the study over time, and were randomized into either the implantable cardiac defibrillator (ICD) arm or the CRT with an ICD (CRT–ICD) arm in a 2:3 ratio. After the trial was completed, it was shown that CRT–ICD reduces the risk of the occurrence of HF or death, especially in patients with a left bundle branch block conduction disturbance (Goldenberg *and others*, 2011; Zareba *and others*, 2011).

Table 2. *Comparison of the empirical coverage probability of CIs of ICERs from* 2000 *simulations and a sample size of* 100

	Uniform cost			Log-normal cost		
Nominal level	New	Reordered	Percentile	New	Reordered	Percentile
0.95	0.938	0.933	0.978	0.934	0.941	0.972
0.90	0.887	0.883	0.949	0.889	0.887	0.948
0.80	0.784	0.788	0.889	0.788	0.779	0.892

New is our proposed method; reordered is the reordered bootstrap percentile method; percentile is the ordinary bootstrap percentile method; the bootstrap replication is 1000.

Table 3. *Estimated mean accumulated costs and life expectancies*, *ICERs and CIs*, *limited to a* 4*-year time horizon*, *for the MADIT-CRT example*

	CRT–ICD	ICD	Difference	95% CI	p-value
Costs ($1000)	62.60	57.05	5.55	1.10, 10.00	0.0146
HF-free YOL	3.29	3.02	0.26	0.12, 0.40	0.0002
Unrestricted YOL	3.61	3.54	0.07	−0.01, 0.15	0.1052

		95% CI of ICER	
ICER($1000/year saved)		New	Bootstrap
Incremental costs for HF-free YOL	21.10	3.40, 64.31	3.38, 62.66
Incremental costs for unrestricted YOL	80.91	—	—

New is our proposed method; bootstrap denotes the bootstrap percentile method with 1000 bootstrap replications.

Owing to the huge costs associated with the implantation of an ICD, a CEA also was conducted based on patients from the US centers, with 503 patients in the ICD arm and 748 in the CRT–ICD arm (Noyes *and others*, 2013). The goal was to evaluate the CE of the CRT–ICD arm when compared with the ICD only arm, restricted to a 4-year horizon, using both the overall survival time and the HF-free survival time as effectiveness measures.

Cost data were collected and available for analysis with start and stop dates for each entry. These were first discounted at a 3% annual rate and then spread out evenly in the interval. These discounted costs were then used to estimate the mean costs within a 4-year time horizon using the ZT estimator (2.1), separately for the CRT–ICD group, and the ICD group. One hundred and twelve (22%) patients in the ICD arm and 65 (9%) patients in the ICD-CRT arm had HFs but kept accumulating costs, with average additional costs of $35 040 and $28 360, respectively. Each patient's survival time was also discounted at a 3% annual rate, and then plugged in the formulae (2.3) or (2.4) to obtain the average unrestricted YOL, or HF-free YOL, within 4 years. In addition, the ICERs comparing the CRT–ICD group and the ICD group were obtained using both the unrestricted and the HF-free YOL. The results are shown in Table 3.

The average health-care expenditures in the CRT–ICD group were higher than the ICD-only group ($62 600 vs $57 050, p-value = 0.0146). The CRT–ICD group also had a larger average HF-Free YOL compared with the ICD group, and the difference (0.26 years) was statistically significant (p-value = 0.0002). These results agreed with the primary study which showed that the CRT–ICD had a significant effect reducing the risk of HF events or death. On the other hand, the difference of unrestricted YOL (0.07 years) was not statistically different (p-value = 0.1052).

Using the method we proposed, the HF-free ICER comparing the CRT–ICD group with the ICD group was $21 100 (95% CI: 3400, 64 310) per one HF-free YOL saved, which is very close to the bootstrap CI (3380, 62 660). Figure 1(a) shows the bootstrap samples on the CE plane, as well as the estimated HF-free ICER and 95% CIs. Since the bootstrap samples lie in the Northeast and Southeast regions of the CE plane, we calculated the bootstrap 95% CI by the bootstrap percentile method. The unrestricted ICER comparing

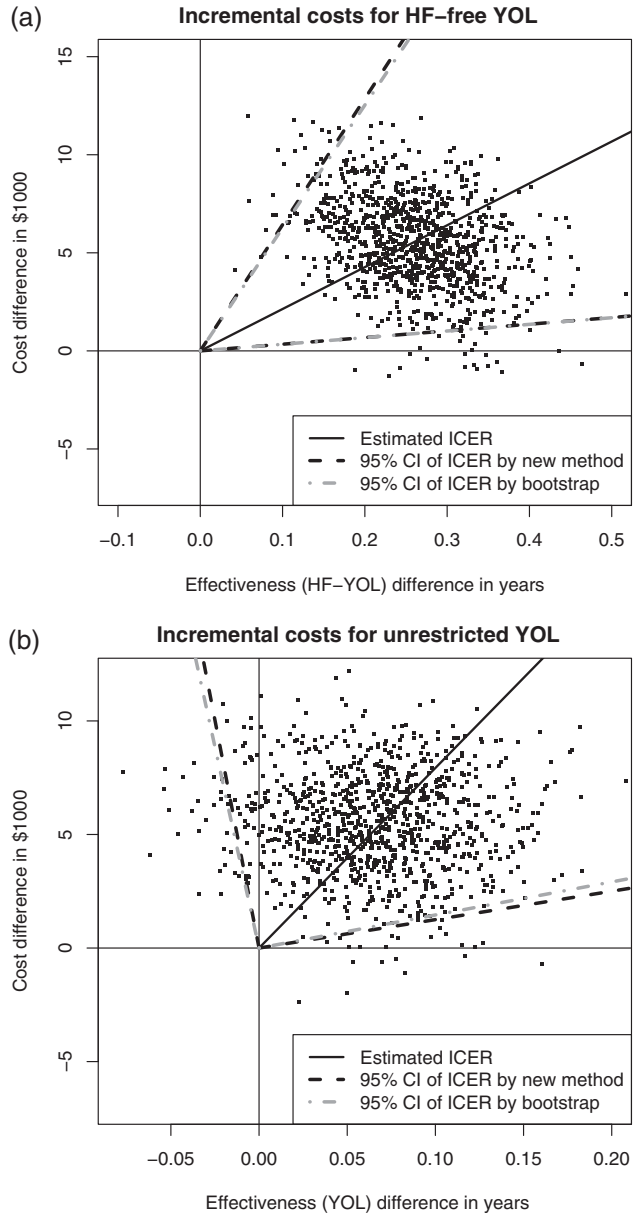

Fig. 1. Estimated 95% CIs of the HF-free ICER (a) and unrestricted ICER (b) for the MADIT-CRT study limited to a 4-year time horizon. The dots are 1000 bootstrap samples. The solid line is the estimated ICER; the dashed lines are CI limits obtained by our method; the gray dot-dashed lines are CI limits obtained by the bootstrap method. The positive y-axis indicates an infinite ICER.

the CRT–ICD group with the ICD group was much higher, \$80 910 per one overall YOL saved. The 95% CI for the unrestricted ICER using both our method and the bootstrap method consists of infinite intervals due to the non-significance of the difference of unrestricted YOL between the two groups, as shown in Figure 1(b).

5. CONCLUSION

In this paper, we consider an important challenge that arises in an actual cost study performed alongside a clinical trial. In a comparison of the CE of a new treatment strategy, the estimates of the effectiveness of the treatment and the costs of the treatment are based on different terminating events. For example, in the MADIT-CRT study, either a HF event or death is used as the terminating event for evaluating the effectiveness of treatment, but death is used for assessing accumulated costs. As in other economic studies conducted in this setting, a censoring problem also complicates the analysis.

We provide a method for estimating consistently the covariance between the cost estimator and the survival estimator under this scenario. This method enables us to construct a correct CI for the ICER. Simulation studies show that our covariance estimator and the CIs perform very well for some practical settings. Our method also accommodates discounting for costs and the survival time, and can easily be extended to obtain ICERs and construct their CIs using quality-adjusted life years as a measure of effectiveness.

The method we propose here expands the usefulness of ICERs in more flexible settings. Further work may be performed on incorporating covariate information to estimate ICERs, and developing software to facilitate the use of these new methods.

SUPPLEMENTARY MATERIAL

Supplementary material is available at http://biostatistics.oxfordjournals.org.

ACKNOWLEDGMENTS

We thank Drs Arthur Moss and Katia Noyes, and Boston Scientific for permitting us to use MADIT-CRT data as an example. We are grateful to Dr W. Jackson Hall and Dr Heejung Bang for their valuable advice and professional help. *Conflict of Interest*: None declared.

FUNDING

This research was supported by R01 HL096575 from the National Heart, Lung, and Blood Institute.

REFERENCES

BANG, H. AND TSIATIS, A. A. (2000). Estimating medical costs with censored data. *Biometrika* **87**, 329–343.

CHAUDHARY, M. A. AND STERNS, S. C. (1996). Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. *Statistics in Medicine* **15**, 1447–1458.

EFRON, B. AND TIBSHIRANI, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Sciences* **1**, 54–77.

EFRON, B. AND TIBSHIRANI, R. (1993). *An Introduction to the Bootstrap*. New York: Chapman and Hall.

FIELLER, E. C. (1954). Some problems in interval estimation. *Journal of the Royal Statistical Society, Series B* **16**, 175–185.

FLEMING, T. R. AND HARRINGTON, D. P. (1991). *Counting Processes and Survival Analysis*. New York: Wiley.

GOLDENBERG, I., HALL, W. J., BECK, C. A., MOSS, A. J., BARSHESHET, A., MCNITT, S., POLONSKY, S., BROWN, M. W. AND ZAREBA, W. (2011). Reduction of the risk of recurring heart failure events with cardiac resynchronization therapy: MADIT-CRT. *Journal of the American College of Cardiology* **58**, 729–737.

HWANG, J. T. G. (1995). Fieller¡s problems and resampling techniques. *Statistica Sinica* **5**, 161–171.

Jiang, G., Wu, J. and Williams, R. G. (2000). Fieller¡s interval and the bootstrap-Fieller interval for the incremental cost-effectiveness ratio. *Health Services and Outcomes Research Methodology* **1**, 291–303.

Jiang, H. and Zhou, X. (2004). Bootstrap confidence intervals for medical costs with censored observations. *Statistics in Medicine* **23**, 3365–3376.

Lin, D. Y., Feuer, E. J., Etzioni, R. and Wax, Y. (1997). Estimating medical costs from incomplete follow-up data. *Biometrics* **53**, 419–434.

McIntosh, E., Barlow, J., Davis, H. and Stewart-Brown, S. (2009). Economic evaluation of an intensive home visiting programme for vulnerable families: a cost-effectiveness analysis of a public health intervention. *Journal of Public Health* **31**, 423–433.

Moss, A. J., Hall, W. J., Cannom, D. S., Klein, H., Brown, M. W., Daubert, J. P., Estes, M., Foster, E., Greenberg, H., Higgins, S. L. *and others*. (2009). Cardiac resynchronization therapy for the prevention of heart failure events. *The New England Journal of Medicine* **361**, 1329–1338.

Mushlin, A. I., Hall, W. J., Zwanziger, J., Gajary, E., Andrews, M., Marron, R., Zou, K. H., Moss, A. J. and for the MADIT Investigators. (1998). The cost-effectiveness of automatic implantable cardiac defibrillators: results from MADIT. *Circulation* **97**, 2129–2135.

Noyes, K., Veazie, P., Hall, W. J., Zhao, H., Buttaccio, A., Thevenet-Morrison, K. and Moss, A. J. (2013). Cost effectiveness of cardiac resynchronization for the prevention of heart failure. *Journal of Cardiovascular Electrophysiology* **24**, 66–74.

O'Brien, B. J. and Briggs, A. H. (2002). Analysis of uncertainty in health care cost-effectiveness studies: an introduction to statistical issues and methods. *Statistical Methods in Medical Research* **11**, 455–468.

O'Hagan, A. and Stevens, J. (2004). On estimators of medical costs with censored data. *Journal of Health Economics* **23**, 615–625.

Pfeifer, P. E. and Bang, H. (2005). Non-parametric estimation of mean customer lifetime value. *Journal of Interactive Marketing* **19**, 48–66.

Raikou, M. and McGuire, A. (2004). Estimating medical care costs under conditions of censoring. *Journal of Health Economics* **23**, 443–479.

Robins, J. and Rotnitzky, A. (1992). Recovery of information and adjustment for dependent censoring using surrogate markers. In: Jewell, N., Dietz, K. and Farewell, V. (editors), *AIDS Epidemiology-Methodological Issues*. Boston: Birkhauser, pp. 297–331.

Robins, J., Rotnitzky, A. and Zhao, L. (1994). Estimation of regression coefficients when some regressors are not always observed. *Journal of the American Statistical Association* **89**, 846–866.

Satten, G. and Datta, S. (2001). The Kaplan–Meier estimator as an inverse-probability-of-censoring weighted average. *The American Statistician* **55**, 207–210.

Wailoo, A. J., Bansback, N., Brennan, A., Michaud, K., Nixon, R. M. and Wolfe, F. (2008). Biologic drugs for rheumatoid arthritis in the medicare program: a cost-effectiveness analysis. *Arthritis & Rheumatism* **58**, 939–946.

Wang, H. and Zhao, H. (2006). Estimating incremental cost-effectiveness ratios and their confidence intervals with differentially censored data. *Biometrics* **62**, 570–575.

Wang, H. and Zhao, H. (2008). A study on confidence intervals for incremental cost-effectiveness. *Biometrical Journal* **50**, 505–514.

Zareba, W., Klein, H., Cygankiewicz, I., Hall, W. J., McNitt, S., Brown, M., Cannom, D., Daubeert, J. P., Eldar, M., Gold, M. R. *and others*. (2011). Effectiveness of cardiac resynchronization therapy by QRS morphology in MADIT-CRT. *Circulation* **123**, 1061–1072.

Zhao, H., Bang, H., Wang, H. and Pfeifer, P. (2007). On the equivalence of some medical cost estimators with censored data. *Statistics in Medicine* **26**, 4520–4530.

ZHAO, H. AND TIAN, L. (2001). On estimating medical cost and incremental cost-effectiveness ratios with censored data. *Biometrics* **57**, 1002–1008.

ZHAO, H. AND WANG, H. (2010). Cost and cost-effectiveness analysis with censored data. In: Faries, D., Leon, A., Haro, J. and Obenchain, R. (editors), *Analysis of Observational Health Care Data Using SAS*. Cary, NC: SAS Press.

ZHAO, H., ZUO, C., CHEN, S. AND BANG, H. (2012). Nonparametric inference for median costs with censored data. *Biometrics* **68**, 717–725.

ZWANZIGER, J., HALL, W. J., DICK, A. W., ZHAO, H., MUSHLIN, A. I., HAHN, R., WANG, H., ANDREWS, M., MOONEY, C., WANG, C. *and others*. (2006). The cost-effectiveness of implantable cardiac defibrillators: results from MADIT II. *Journal of the American College of Cardiology* **47**, 2310–2318.

[*Received July 11, 2012; revised December 6, 2012; accepted for publication January 16, 2013*]