

## Incremental net benefit in randomized clinical trials with quality-adjusted survival

Andrew R. Willan<sup>1,2,\*</sup>, Eric Bingshu Chen<sup>3</sup>, Richard J. Cook<sup>3</sup> and D. Y. Lin<sup>4</sup>

<sup>1</sup>*Program in Population Health Sciences, Research Centre, Hospital for Sick Children,  
555 University Avenue, Toronto ON, M5G 1X8, Canada*

<sup>2</sup>*Department of Public Health Sciences, University of Toronto, Canada*

<sup>3</sup>*Department of Statistics & Actuarial Science, 200 University Avenue West, University of Waterloo,  
Waterloo, ON, N2L 3G1, Canada*

<sup>4</sup>*Department of Biostatistics, University of North Carolina, 3101E McGavran-Greenberg Hall, CB #7420  
Chapel Hill, NC 27599-7420, U.S.A.*

### SUMMARY

Owing to induced dependent censoring, estimating mean costs and quality-adjusted survival in a cost-effectiveness comparison of two groups using standard life-table methods leads to biased results. In this paper we propose methods for estimating the difference in mean costs and the difference in mean effectiveness, together with their respective variances and covariance in the presence of dependent censoring. We consider the situation in which the measure of effectiveness is either the probability of surviving a duration of interest or mean quality-adjusted survival time over a duration of interest. The methods are illustrated in an example using an incremental net benefit analysis. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS: incremental net benefit; quality-adjusted survival; cost-effectiveness

### 1. INTRODUCTION

It is becoming increasingly common for researchers to collect patient-level cost data in addition to effectiveness outcomes in randomized clinical trials. As a result there have been many publications [1–27] regarding the development of statistical methods for the design and analysis of cost-effectiveness studies. Initially authors concentrated on providing inference on incremental cost-effectiveness ratios (ICER), but emphasis has recently shifted to incremental net benefit (INB) [21–27], for which the ICER can be thought of as a special case [25–27]. To conduct an INB analysis one need only estimate the difference in mean effectiveness and the difference in mean costs, along with their associated variances and covariance. If all patients are followed for the duration of interest, sample means, variances and covariances are all that are required [26, 27]. However, for censored data, calculations become considerably

\* Correspondence to: Andrew R. Willan, 555 University Avenue, Toronto, ON, M5G 1X8, Canada.

† E-mail: willana@mcmaster.ca

more complex [27–35]. Willan and Lin [27] provide solutions for parameter estimation when censoring is present and the measure of effectiveness is mean survival. Recently [29–35] other methods have been proposed for estimating mean cost and quality-adjusted survival in the presence of censoring. Simulation results [31–33] indicated that these estimators have similar bias and variance as those proposed by Lin *et al.* [28]. In this paper we provide a solution, using methods similar to those of Lin *et al.* [28] and Willan and Lin [27], for the situation in which censoring is present and the measure of effectiveness is either the probability of surviving a duration of interest or mean quality-adjusted survival time over a duration of interest. The statistical model and the solutions are given in Section 2, followed by examples of INB analyses in Section 3.

## 2. METHODS

### 2.1. Notation and derivation

In a two-arm randomized controlled trial let  $c_{ji}$  be the total cost over the duration of interest for patient  $i$  on therapy  $j$ , where  $j = T$  (treatment),  $S$  (standard),  $i = 1, 2, \dots, n_j$ , and  $n_j$  is the number of patients randomized to therapy  $j$ . Let  $E(C_{ji}) = v_j$  and  $\Delta_c = v_T - v_S$ . For patient  $i$  on therapy  $j$ , let  $D_{ji}$  and  $U_{ji}$  be the times from randomization to death and censoring, respectively. Let  $S_j(t) = \Pr(D_{ji} \geq t)$ ,  $X_{ji} = \min(D_{ji}, U_{ji})$  and  $\delta_{ji} = I\{D_{ji} < U_{ji}\}$ , where  $I$  is the indicator function.

A valid method for estimating  $\Delta_c$  is proposed by Lin *et al.* [28]. In this method the duration of interest  $(0, \tau]$  is divided into  $K$  intervals  $[a_k, a_{k+1})$ , where  $0 = a_1 < a_2 < \dots < a_{K+1} = \tau$ . Let  $C_{jki}$  be the observed cost for patient  $i$  in group  $j$  during interval  $k$ . Let  $Y_{jki} = I\{x_{ji} \geq a_k \text{ and } (X_{ji} \geq a_{k+1} \text{ or } \delta_{ji} = 1)\}$  indicate whether patient  $i$  on therapy  $j$  is known to be alive at  $a_k$  and is not censored in  $[a_k, a_{k+1})$ , and let  $Y_{jk} = \sum_{i=1}^{n_j} Y_{jki}$ , and let  $\bar{C}_{jk} = \sum_{i=1}^{n_j} (Y_{jki} C_{jki}) / Y_{jk}$  be the average cost of these patients. Then  $v_j$  is estimated by  $\hat{v}_j = \sum_{k=1}^K \hat{S}_j(a_k) \bar{C}_{jk}$ , where  $\hat{S}_j(t)$  is the product-limit estimate for  $S_j(t)$ , the probability of surviving beyond  $t$  in the  $j$ th group. The estimator of the variance of  $\hat{v}_j$  is given by

$$\hat{V}(\hat{v}_j) = \sum_{i=1}^{n_j} \sum_{k=1}^K \sum_{l=1}^K \hat{Z}_{jki}^{(c)} \hat{Z}_{jli}^{(c)}$$

where

$$\hat{Z}_{jki}^{(c)} = \frac{(C_{jki} - \bar{C}_{jk}) \hat{S}_j(a_k) Y_{jki}}{Y_{jk}} - \hat{S}_j(a_k) \bar{C}_{jk} \left( \frac{I\{X_{ji} \leq a_k\} \delta_{ji}}{R_{ji}} - \sum_{l=1}^{n_j} \frac{I\{X_{jl} \leq \min(X_{ji}, a_k)\} \delta_{jl}}{R_{jl}^2} \right)$$

and  $R_{ji} = \sum_{l=1}^{n_j} I\{X_{jl} \geq X_{ji}\}$ . Thus  $\Delta_c$  is estimated by  $\hat{\Delta}_c = \hat{v}_T - \hat{v}_S$  and the estimated variance for  $\hat{\Delta}_c$  is given by  $\hat{V}(\hat{\Delta}_c) = \hat{V}(\hat{v}_T) + \hat{V}(\hat{v}_S)$ . Details are given in the Appendix.

### 2.2. Probability of surviving the duration of interest

If the measure of effectiveness is the probability of surviving the duration of interest, then the parameter of interest in the comparison of therapies is  $\Delta_e = S_T(\tau) - S_S(\tau)$ . Let  $t_{jh}$ , where

$t_{j,h-1} < t_{jh}$ ,  $h = 1, 2, \dots, L_j$ , be unique times to death for patients on therapy  $j$ . Then the product-limit estimator of  $S_j(\tau)$  is given by

$$\hat{S}_j(\tau) = \prod_{h|t_{jh} < \tau} (1 - d_{jh}/n_{jh})$$

where  $d_{jh}$  are the number of deaths on therapy  $j$  at time  $t_{jh}$ , and  $n_{jh}$  is the number of patients still at risk at time  $t_{jh}$ . The estimator of the variance of  $\hat{S}_j(\tau)$  is given by  $\hat{V}(\hat{S}_j(\tau)) = \sum_{i=1}^{n_j} (\hat{Z}_{ji}^{(e)})^2$ , where

$$\hat{Z}_{ji}^{(e)} = -\hat{S}_j(\tau) \left( \frac{I\{X_{ji} \leq \tau\} \delta_{ji}}{R_{ji}} - \sum_{l=1}^{n_j} \frac{I\{X_{jl} \leq \min(\tau, X_{ji})\} \delta_{jl}}{R_{jl}^2} \right)$$

The estimator of  $\Delta_e$  is given by  $\hat{\Delta}_e = \hat{S}_T(\tau) - \hat{S}_S(\tau)$ , with estimated variance given by  $\hat{V}(\hat{\Delta}_e) = \hat{V}(\hat{S}_T(\tau)) + \hat{V}(\hat{S}_S(\tau))$ . The estimated covariance between  $\hat{\Delta}_e$  and  $\hat{\Delta}_c$  is given by

$$\begin{aligned} \hat{C}(\hat{\Delta}_e, \hat{\Delta}_c) &= \hat{C}(\hat{S}_T(\tau), \hat{v}_T) + \hat{C}(\hat{S}_S(\tau), \hat{v}_S) \\ &= \sum_{j=T,S} \sum_{i=1}^{n_j} \left( \hat{Z}_{ji}^{(e)} \sum_{k=1}^K \hat{Z}_{jki}^{(c)} \right) \end{aligned}$$

Details are given in the Appendix.

### 2.3. Quality-adjusted survival

The standard application of survival analysis methods to quality-adjusted survival leads to biased estimators of the mean quality-adjusted survival [27, 29, 36]. Consequently, we propose using methods similar to those used in Section 2.1. Let  $q_{ji}$  be the quality adjusted survival for the duration of interest for patient  $i$  on therapy  $j$ . Let  $E(q_{ji}) = \mu_j$  and  $\Delta_e = \mu_T - \mu_S$ . Recall that the duration of interest  $(0, \tau)$  is divided into the  $K$  sub-intervals  $[a_k, a_{k+1})$ , where  $0 = a_1 < a_2 < \dots < a_{K+1} = \tau$ . Let  $q_{jki}$  be the observed quality of life experienced by the patient  $i$  on therapy  $j$  during time interval  $k$ . The quantities  $q_{jki}$  can be determined in the following way. Suppose patient  $i$  on therapy  $j$  has  $B_{ji}$  quality of life measures taken at  $t_{ji1}, t_{ji2}, \dots, t_{jim_{ji}}$  with respective scores  $Q_{ji1}, Q_{ji2}, \dots, Q_{jim_{ji}}$ . Then  $q_{jki} = \int_{a_k}^{a_{k+1}} Q(t) dt$ , where

$$Q(t) = \begin{cases} Q_{ji1}: & 0 \leq t < t_{ji1} \\ Q_{jih} + \frac{(Q_{ji,h+1} - Q_{jih})(t - t_{jih})}{t_{ji,h+1} - t_{jih}}: & t_{jih} \leq t < t_{ji,h+1} \\ Q_{jim_{ji}}: & t_{jim_{ji}} \leq t < X_{ji} \\ 0: & t \geq X_{ji} \end{cases}$$

Let  $\bar{q}_{jk} = \sum_{i=1}^{n_j} (Y_{jki} q_{jki}) / Y_{jk}$ . Then  $\hat{\mu}_j = \sum_{k=1}^K \hat{S}_j(a_k) \bar{q}_{jk}$ . The estimator of the variance of  $\hat{\mu}_j$  is given by

$$\hat{V}(\hat{\mu}_j) = \sum_{i=1}^{n_j} \sum_{k=1}^K \sum_{l=1}^K \hat{Z}_{jki}^{(e)} \hat{Z}_{jli}^{(e)}$$

where

$$\hat{Z}_{jki}^{(e)} = \frac{(q_{jki} - \bar{q}_{jk})\hat{S}_j(a_k)Y_{jki}}{Y_{jk}} - \hat{S}_j(a_k)\bar{q}_{jk} \left( \frac{I\{X_{ji} \leq a_k\}\delta_{ji}}{R_{ji}} - \sum_{l=1}^{n_j} \frac{I\{X_{jl} \leq \min(a_k, X_{ji})\}\delta_{jl}}{R_{jl}^2} \right)$$

Thus  $\hat{\Delta}_e = \hat{\mu}_T - \hat{\mu}_S$  and  $\hat{V}(\hat{\Delta}_e) = \hat{V}(\hat{\mu}_T) + \hat{V}(\hat{\mu}_S)$ . The estimated covariance between  $\hat{\Delta}_e$  and  $\hat{\Delta}_c$  is given by

$$\hat{C}(\hat{\Delta}_e, \hat{\Delta}_c) = \sum_{j=T,S} \sum_{i=1}^{n_j} \sum_{k=1}^K \sum_{l=1}^K \hat{Z}_{jki}^{(e)} \hat{Z}_{jli}^{(c)}$$

Details are given in the Appendix. An SAS program for estimating the parameters for all three measures of effectiveness can be found on the web site [www.andywillan.com/programs](http://www.andywillan.com/programs).

### 3. EXAMPLES

#### 3.1. Incremental net benefit analysis

The INB, given as a function of the willingness-to-pay  $\lambda$ , is defined as  $b(\lambda) = \lambda\Delta_e - \Delta_c$  and is estimated by  $\hat{b}(\lambda) = \lambda\hat{\Delta}_e - \hat{\Delta}_c$ . The variance of  $\hat{b}(\lambda)$  is estimated by  $\hat{V}(\hat{b}(\lambda)) = \lambda^2\hat{V}(\hat{\Delta}_e) + \hat{V}(\hat{\Delta}_c) - 2\lambda\hat{C}(\hat{\Delta}_e, \hat{\Delta}_c)$ . Thus the null hypothesis,  $H_0 : b(\lambda) = 0$ , versus the alternative hypothesis  $H_1 : b(\lambda) > 0$ , can be rejected at the level  $\alpha$  if  $\hat{b}(\lambda)/\sqrt{\{V(\hat{b}(\lambda))\}}$  exceeds  $z_{1-\alpha}$ , where  $z_{1-\alpha}$  is the  $1 - \alpha$  percentile of the standard normal random variable. In addition, the  $100(1 - \alpha/2)$  per cent confidence limits for  $b(\lambda)$  are given by  $\hat{b}(\lambda) \pm z_{1-\alpha/2}\sqrt{\{V(\hat{b}(\lambda))\}}$ . Most often an INB analysis is performed as a function of  $\lambda$  and illustrated in a plot of  $\hat{b}(\lambda)$  versus  $\lambda$ . The slope of the plot is  $\hat{\Delta}_e$  and the vertical intercept is  $-\hat{\Delta}_c$ . It is worth noting that the horizontal intercept of  $\hat{b}(\lambda)$  is the estimate of the ICER, that is,  $\hat{b}(\hat{\Delta}_c/\hat{\Delta}_e) = 0$ . Furthermore, the horizontal intercepts of the confidence limits for  $b(\lambda)$  provide the Fieller confidence limits for the ICER [25–27].

Therefore, to carry out a statistical analysis of INB the following five parameters must be estimated:  $\Delta_e$ ;  $\Delta_c$ ;  $V(\hat{\Delta}_e)$ ;  $V(\hat{\Delta}_c)$ ;  $C(\hat{\Delta}_e, \hat{\Delta}_c)$ . In Section 2, estimators for these parameters are derived for the situation in which the measure of effectiveness is either the probability of surviving the duration of interest or mean quality-adjusted survival. These methods are illustrated in Sections 3.2 and 3.3, respectively.

#### 3.2. Probability of surviving the duration of interest

In a trial of patients at risk of cardiac arrest, a total of 659 patients with resuscitated ventricular fibrillation or sustained ventricular tachycardia or with unmonitored syncope were randomized between amiodarone (S) and implantable cardioverter defibrillator (T) during the period from October 1990 to January 1997. Owing to budgetary constraints, the costs were collected on the first 430 patients only. The primary outcome measure was all-cause mortality. The clinical

Table I. Sample sizes and parameter estimates for the example with probability surviving 77 months as the measure of effectiveness.

	$n_j$	$\hat{S}_j(\tau)$	$\hat{v}_j$	$\hat{V}(\hat{S}_j(\tau))$	$\hat{V}(\hat{v}_j)$	$\hat{C}(\hat{S}_j(\tau), \hat{v}_j)$
Treatment	212	0.587	87 103	0.00284	8 461 538	13.560
Standard	218	0.567	38 859	0.00196	6 517 608	-5.592

$$\hat{\Delta}_e = 0.0207; \hat{\Delta}_c = 48\,244; \hat{V}(\hat{\Delta}_e) = 0.00481; \hat{V}(\hat{\Delta}_c) = 14\,979\,146; \hat{C}(\hat{\Delta}_e, \hat{\Delta}_c) = 7.968.$$

results are reported in Connolly *et al.* [37], and the economic evaluation in O'Brien *et al.* [38]. A non-significant reduction in the risk of death was observed with T, from 10.2 per cent per year to 8.3 per cent per year (19.7 per cent relative risk reduction; 95 per cent confidence interval, -7.7 to 40 per cent;  $P = 0.142$ ). For the economic results reported below, the duration of interest is 6.42 (77 months), that is,  $\tau = 6.42$ . Censoring is present since not all patients were followed for 77 months.

If the measure of effectiveness is the probability of surviving, then the willingness-to-pay is expressed as CAD\$ per life saved. The relevant parameters, estimated using the procedures given in Section 2.2, can be found in Table I. Cost is given in CAD\$.  $\hat{b}(\lambda)$  and the corresponding 95 per cent confidence limits are plotted as a function of  $\lambda$  in Figure 1. For  $\lambda = \text{CAD\$ } 100\,000$  per life saved, the estimate of INB is -CAD\$ 46 174 with confidence interval -CAD\$ 61 540 to -CAD\$ 30 808. The slope of  $\hat{b}(\lambda)$  is positive (that is,  $\hat{\Delta}_e = 0.0207 > 0$ ), illustrating that treatment was observed to increase the probability of surviving 77 months. The estimate of the ICER, given by the horizontal intercept, is 2 330 898 CAD\$/life saved. The Fieller lower confidence limit for the ICER, given by the horizontal intercept of the upper confidence limit for  $b(\lambda)$ , is 305 094 CAD\$/life saved. The lower confidence limit for  $b(\lambda)$  does not cross the horizontal axis, indicating that there is no upper limit for the ICER. One interprets this to mean that arbitrarily large values for the ICER are not inconsistent with the data, which is to be expected since the difference in effectiveness is not statistically significant.

### 3.3. Quality-adjusted survival

Utility-based quality of life data were not collected in this trial. Therefore, for illustration we randomly generated the quality of life scores  $q_{jki}$  as  $(a_{k+1} - a_k)\Phi(W_{jki})$ , where  $W_{jki}$  is a randomly generated normally distributed random variable with mean  $(1 + 1.75 \times I\{j = T\})$  and unit variance and  $\Phi$  is the cumulative probability distribution for the standard normal random variable, that is,  $\Phi(z_{1-\alpha}) = 1 - \alpha$ . Thereby, we have made treatment artificially superior to standard with respect to quality-adjusted survival.

The relevant parameters, estimated using the procedures given in Section 2.3, can be found in Table II. Because the measure of effectiveness is quality-adjusted survival, the willingness-to-pay is expressed as CAD\$ per quality-adjusted life-year (QALY).  $\hat{b}(\lambda)$  and the corresponding 95 per cent confidence limits are plotted as a function of  $\lambda$  in Figure 2. For  $\lambda = \text{CAD\$ } 50\,000$  per QALY, the estimate of INB is CAD\$ 10 101 with confidence interval -CAD\$ 9162 to CAD\$ 29 364. The slope of  $\hat{b}(\lambda)$  is positive (that is,  $\hat{\Delta}_e = 1.167 > 0$ ), illustrating that treatment was observed to increase quality-adjusted survival over the 77-month duration of interest. The vertical intercept is negative (that is,  $\hat{\Delta}_c = 48\,244 > 0$ ), illustrating that treatment

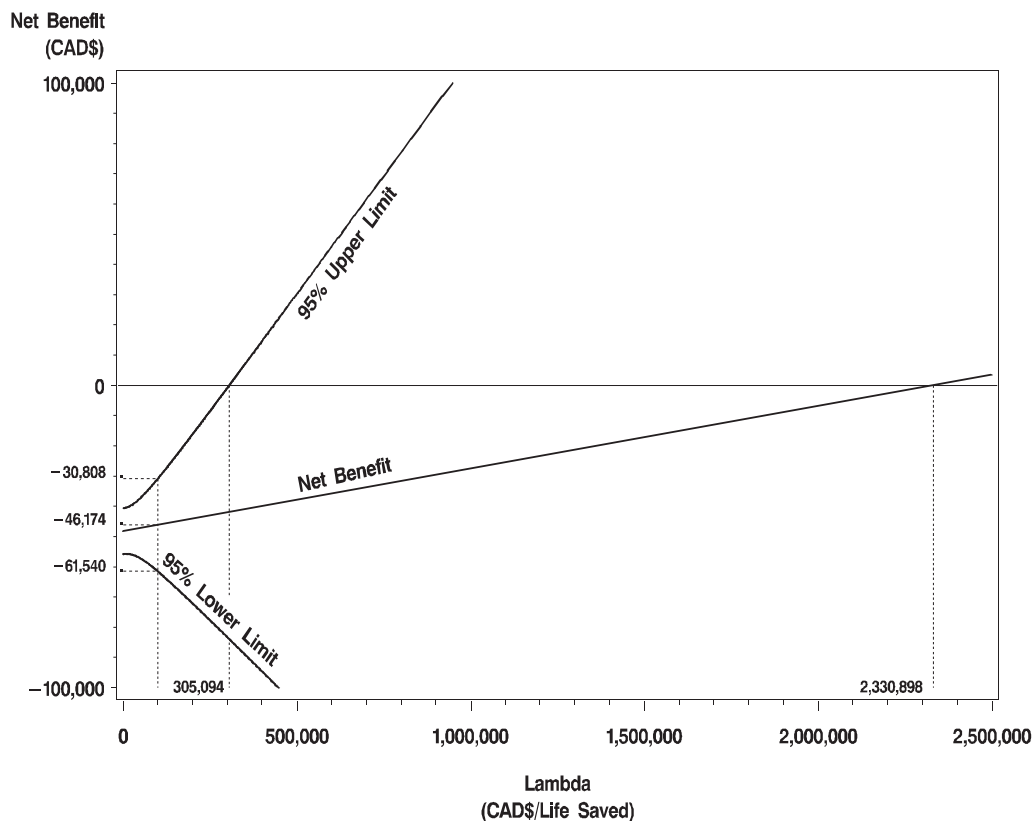


Figure 1. Incremental net benefit and confidence limits for probability of surviving duration of interest.

Table II. Sample sizes and parameter estimates for the example with quality-adjusted survival as the measure of effectiveness.

	$n_j$	$\hat{\mu}_j$	$\hat{v}_j$	$\hat{V}(\hat{\mu}_j)$	$\hat{V}(\hat{v}_j)$	$\hat{C}(\hat{\mu}_j, \hat{v}_j)$
Treatment	212	4.860	87 103	0.02315	8 461 538	122.898
Standard	218	3.693	38 859	0.01471	6 517 608	7.351

$$\hat{\Delta}_c = 1.167; \hat{\Delta}_c = 48\,244; \hat{V}(\hat{\Delta}_c) = 0.03786; \hat{V}(\hat{\Delta}_c) = 14\,979\,146; \hat{C}(\hat{\Delta}_c, \hat{\Delta}_c) = 130.25.$$

was observed to increase cost. The graphs provide a confidence interval for  $\Delta_c$  of CAD\$ 40 658 to CAD\$ 55 830 (the negative of the vertical intercepts). The estimate of the ICER, given by the horizontal intercept, is 41 344 CAD\$/QALY. The Fieller confidence interval for the ICER, defined by the horizontal intercept of the confidence limits for  $b(\lambda)$ , is 30 442 to 61 310 CAD\$/QALY.

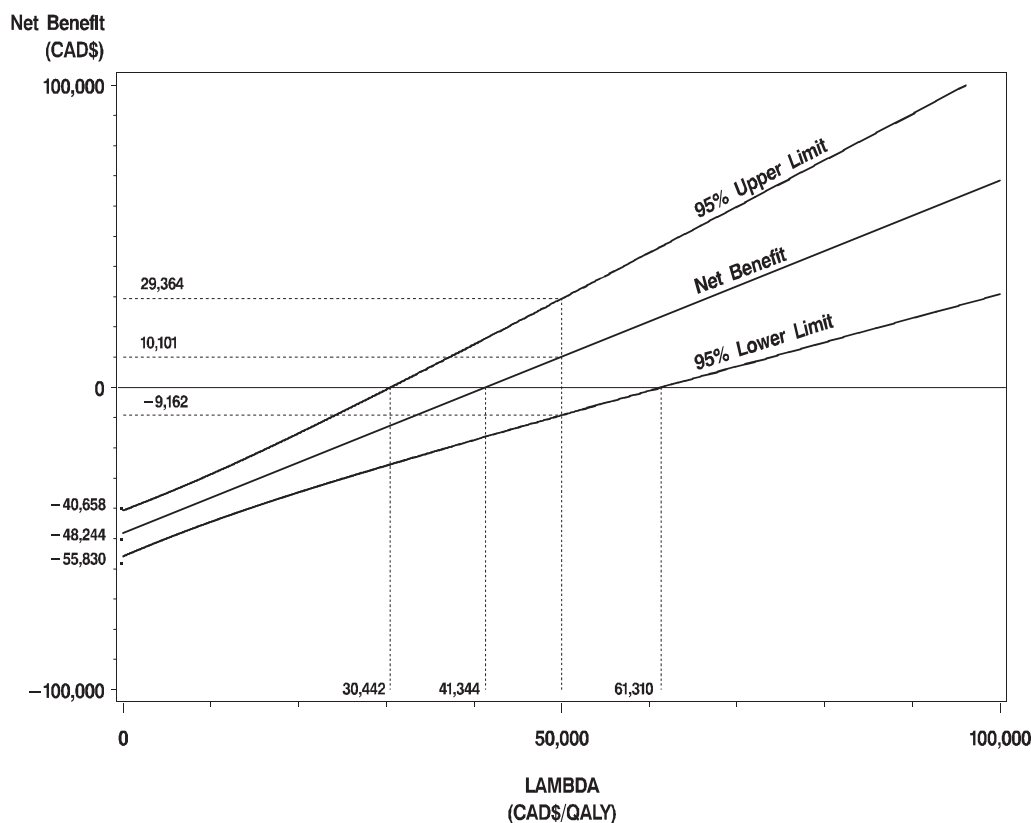


Figure 2. Incremental net benefit and confidence limits for mean quality-adjusted survival.

#### 4. DISCUSSION

In this paper we have extended the methods for performing cost-effectiveness analysis in the presence of censoring proposed by Willan and Lin [27] to the situation in which the measure of effectiveness is either the probability of surviving the duration of interest or mean quality-adjusted survival. We have done this in the context of INB, but by varying the willingness-to-pay in a sensitivity analysis, inference about the ICER is also possible.

Methods initially proposed by Lin *et al.* [28], often referred to as the *direct* method, were used to account for the induced dependent censoring present on the cost and quality-adjusted survival scales. Since patients accumulated cost and quality-adjusted survival at different rates, there is a positive correlation between the amounts accumulated to death and censoring. Because of the positive correlation, the application of standard life-table methods would provide overestimates of mean cost and quality-adjusted survival. Other methods, sometimes referred to as *inverse weighting*, have been proposed for estimating mean costs [30, 31, 33, 34] and mean quality-adjusted survival [29] in the presence of censoring. The advantage of *inverse weighting* is that it allows for arbitrary censoring, whereas for completely unbiased

estimators, the *direct* methods requires that the censoring take place at the beginning or end of the intervals defined by the  $a_k$ 's.

#### APPENDIX

According to Lin *et al.* [28],  $n_j^{1/2}(\hat{v}_j - v_j)$  can be represented by a sum of zero-mean i.i.d. random variables:

$$n_j^{1/2}(\hat{v}_j - v_j) = n_j^{-1/2} \sum_{i=1}^{n_j} \sum_{k=1}^K Z_{jki}^{(c)} + o_p(1)$$

where

$$Z_{jki}^{(c)} = \frac{S_j(a_k) Y_{jki} (C_{jki} - v_{jk})}{E(Y_{jki})} - v_{jk} S_j(a_k) \int_0^{a_k} \frac{dM_{ji}(t)}{\Pr(X_{ji} \geq t)}$$

$$v_{jk} = E(C_{jki} | Y_{jki} = 1)$$

$$M_{ji}(t) = \delta_{ji} I\{X_{ji} \leq t\} - \int_0^t I\{X_{ji} \geq t\} d\Lambda_j(t)$$

and  $\Lambda_j(t)$  is the cumulative hazard function of the survival time in the  $j$ th group, where  $E$  is the expectation function. Similarly,  $n_j^{1/2}(\hat{S}_j(\tau) - S_j(\tau))$  and  $n_j^{1/2}(\hat{\mu}_j - \mu)$  can be represented by sums of zero-mean i.i.d. random variables:

$$n_j^{1/2}(\hat{S}_j(\tau) - S_j(\tau)) = n_j^{-1/2} \sum_{i=1}^{n_j} Z_{ji}^{(e)} + o_p(1)$$

$$n_j^{1/2}(\hat{\mu}_j - \mu_j) = n_j^{-1/2} \sum_{i=1}^{n_j} \sum_{k=1}^K Z_{jki}^{(e)} + o_p(1)$$

where

$$Z_{ji}^{(e)} = -S_j(\tau) \int_0^\tau \frac{dM_{ji}(t)}{\Pr(X_{ji} \geq t)}$$

$$Z_{jki}^{(e)} = \frac{S_j(a_k) Y_{jki} (\mu_{jki} - \mu_{jk})}{E(Y_{jki})} - S_j(a_k) \mu_{jk} \int_0^{a_k} \frac{dM_{ji}(t)}{\Pr(X_{ji} \geq t)}$$

and  $\mu_{jk} = E(q_{jki} | Y_{jki} = 1)$ . It then follows from the multivariate central limit theorem that  $n_j^{1/2}(\hat{v}_j - v_j)$  and  $n_j^{1/2}(\hat{S}_j(\tau) - S_j(\tau))$  are asymptotically bivariate normal with zero means

$$V(n_j^{1/2}(\hat{v}_j - v_j)) = E \left( \sum_{k=1}^K \sum_{l=1}^K Z_{jk1}^{(c)} Z_{jl1}^{(c)} \right)$$

$$V(n_j^{1/2}(\hat{S}_j(\tau) - S_j(\tau))) = E(Z_{j1}^{(e)})^2$$

$$C(n_j^{1/2}(\hat{S}_j(\tau) - S_j(\tau)), n_j^{1/2}(\hat{v}_j - v_j)) = E \left( Z_{j1}^{(e)} \sum_{k=1}^K Z_{jk1}^{(c)} \right)$$



and  $n_j^{1/2}(\hat{v}_j - v_j)$  and  $n_j^{1/2}(\hat{\mu}_j - \mu_j)$  are also asymptotically bivariate normal with zero means

$$V(n_j^{1/2}(\hat{\mu}_j - \mu_j)) = E \left( \sum_{k=1}^K \sum_{l=1}^K Z_{jk1}^{(e)} Z_{jl1}^{(e)} \right)$$

and

$$C(n_j^{1/2}(\hat{\mu}_j - \mu_j), n_j^{1/2}(\hat{v}_j - v_j)) = E \left( \sum_{k=1}^K \sum_{l=1}^K Z_{jk1}^{(e)} Z_{jl1}^{(c)} \right)$$

where  $V$  and  $C$  are the variance and covariance functions, respectively. For large samples,  $V(\hat{v}_j)$ ,  $V(\hat{S}_j(\tau))$ ,  $V(\hat{\mu}_j)$ ,  $C(\hat{S}_j(\tau), \hat{v}_j)$  and  $C(\hat{\mu}_j, \hat{v}_j)$  are estimated, respectively, by

$$\begin{aligned} \hat{V}(\hat{v}_j) &= \sum_{i=1}^{n_j} \sum_{k=1}^K \sum_{l=1}^K \hat{Z}_{jki}^{(c)} \hat{Z}_{jli}^{(c)} \\ \hat{V}(\hat{S}_j(\tau)) &= \sum_{i=1}^{n_j} (\hat{Z}_{ji}^{(e)})^2 \\ \hat{V}(\hat{\mu}_j) &= \sum_{i=1}^{n_j} \sum_{k=1}^K \sum_{l=1}^K \hat{Z}_{jki}^{(e)} \hat{Z}_{jli}^{(e)} \\ \hat{C}(\hat{S}_j(\tau), \hat{v}_j) &= \sum_{i=1}^{n_j} \left( \hat{Z}_{ji}^{(e)} \sum_{k=1}^K \hat{Z}_{jki}^{(c)} \right) \end{aligned}$$

and

$$\hat{C}(\hat{\mu}_j, \hat{v}_j) = \sum_{i=1}^{n_j} \sum_{k=1}^K \sum_{l=1}^K \hat{Z}_{jki}^{(e)} \hat{Z}_{jli}^{(c)}$$

where  $\hat{Z}_{jki}^{(c)}$ ,  $\hat{Z}_{ji}^{(e)}$  and  $\hat{Z}_{jki}^{(e)}$  are obtained from  $Z_{jki}^{(c)}$ ,  $Z_{ji}^{(e)}$  and  $Z_{jki}^{(e)}$ , respectively, by replacing all the unknown parameters with their sample estimators.

ACKNOWLEDGEMENTS

The authors wish to thank Dr Gary Foster for producing the figures and Ms Susan Tomlinson for her meticulous proof reading.

REFERENCES

1. O'Brien BJ, Drummond MF, Labelle RJ, Willan AR. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. *Medical Care* 1994; **32**:150–163.
2. Mullahy J, Manning W. Statistical issues of cost-effectiveness analysis. In *Valuing Health Care*, Sloan F (ed.). Cambridge University Press: Cambridge, 1994.
3. van Hout BA, Al MJ, Gordon GS, Rutten FFH. Costs, effects and C/E ratios alongside a clinical trial. *Health Economics* 1994; **3**:309–319.
4. Wakker P, Klaassen MP. Confidence intervals for cost/effectiveness ratios. *Health Economics* 1995; **4**: 373–381.
5. Willan AR, O'Brien BJ. Confidence intervals for cost-effectiveness ratios: an application of Fieller's theorem. *Health Economics* 1996; **5**:297–305.
6. Chaudhary MA, Stearns SC. Confidence intervals for cost-effectiveness ratios: an example from a randomized trial. *Statistics in Medicine* 1996; **15**:1447–1458.
7. Mullahy J. What you don't know can't hurt you? Statistical issues and standards for medical technology evaluation. *Medical Care* 1996; **34**(12 Suppl):DS124–DS135.

8. Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost effectiveness analysis. In *Cost Effectiveness in Health and Medicine*, Gold MR, Siegel JE, Russell LB, Weinstein MC (eds). Oxford University Press: New York, 1996.
9. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps; a non-parametric approach to confidence interval estimation. *Health Economics* 1997; **6**:327–340.
10. Polsky D, Glick HA, Willke R, Schulman K. Confidence intervals for cost-effectiveness ratios: a comparison of four methods. *Health Economics* 1997; **6**:243–252.
11. Severens JL, De Boo TM, Konst EM. Uncertainty of incremental cost-effectiveness ratios. *International Journal of Technology Assessment in Health Care* 1999; **15**:608–614.
12. Briggs AH, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. *Health Economics* 1998; **7**:723–740.
13. Briggs AH, Gray AM. Power and sample size calculations for cost-effectiveness studies. *Medical Decision Making* 1998; **18**(Suppl):S81–S92.
14. Laska EM, Meisner M, Siegel C. Power and sample in cost-effectiveness analysis. *Medical Decision Making* 1999; **19**:339–343.
15. Briggs AH, Mooney CZ, Wonderling DE. Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using monte carlo simulation. *Statistics in Medicine* 1999; **18**:3245–3262.
16. Briggs AH, Gray AM. Handling uncertainty in economic evaluations of healthcare interventions. *British Medical Journal* 1999; **319**:635–638.
17. Heitjan DF, Moskowitz AJ, Whang W. Bayesian estimation of cost-effectiveness ratios from clinical trials. *Health Economics* 1999; **8**:191–201.
18. Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis. *Health Economics* 1999; **8**:257–261.
19. Willan AR, O'Brien BJ. Sample size and power issues in estimating incremental cost-effectiveness ratios from clinical trials data. *Health Economics* 1999; **8**:203–211.
20. Gardiner JC, Huebner M, Jetton J, Bradley CJ. Power and sample size assessments for test of hypotheses on cost-effectiveness ratios. *Health Economics* 2000; **9**:227–234.
21. Phelps CE, Mushlin AI. On the (near) equivalence of cost-effectiveness and cost-benefit analysis. *International Journal of Technology Assessment in Health Care* 1991; **7**:12–21.
22. Ament A, Baltussen R. The interpretation of results of economic evaluation: explicating the value of health. *Health Economics* 1997; **6**:625–635.
23. Stinnett AA, Mallahy J. Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical Decision Making* 1998; **18**(Suppl):S68–S80.
24. Tambour M, Zethraeus N, Johannesson M. A note on confidence intervals in cost-effectiveness analysis. *International Journal of Technology Assessment* 1998; **14**:467–471.
25. Heitjan DF. Fieller's method and net health benefit. *Health Economics* 2000; **9**:327–335.
26. Willan AR. Analysis, sample size and power for estimating incremental net benefit from clinical trials data. *Controlled Clinical Trials* 2001; **22**:228–237.
27. Willan AR, Lin DY. Incremental net benefit in randomized clinical trials. *Statistics in Medicine* 2001; **20**:1563–1574.
28. Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. *Biometrics* 1997; **53**:419–434.
29. Zhao H, Tsiatis AA. A consistent estimator for the distribution of quality adjusted survival time. *Biometrika* 1997; **84**:339–348.
30. Lin DY. Linear regression analysis of censored medical costs. *Biostatistics* 2000; **1**:35–47.
31. Bang H, Tsiatis AA. Estimating medical costs with censored data. *Biometrika* 2000; **87**:329–343.
32. Carides GW, Heyse JF, Iglewicz B. A regression-based method for estimating mean treatment cost in the presence of right-censoring. *Biostatistics* 2000; **1**:229–313.
33. Strawderman RL. Estimating the mean of an increasing stochastic process at a censored stopping time. *Journal of the American Statistical Association* 2000; **95**:1192–1208.
34. Zhao H, Tsiatis AA. Efficient estimation of the distribution of quality adjusted survival time. *Biometrics* 1999; **55**:1101–1107.
35. Zhao H, Tsiatis AA. Testing equality of survival functions of quality adjusted lifetime. *Biometrics* 2001; **57**:861–867.
36. Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. *Statistics in Medicine* 1990; **9**:1259–1276.
37. Connolly SJ, Gent M, Roberts RS, Dorian P, Roy D, Sheldon RS, Mitchell LB, Green MS, Klein GJ, O'Brien B. Canadian implantable defibrillator study (CIDS): a randomized trial of the implantable cardioverter defibrillator against amiodarone. *Circulation* 2000; **101**:1297–1302.
38. O'Brien BJ, Connolly SJ, Goeree R, Blackhouse G, Willan AR, Yee R, Roberts RS, Gent M. Cost-effectiveness of the implantable cardioverter defibrillator results from the Canadian implantable defibrillator study (CIDS). *Circulation* 2000; **103**:1416–1421.