Blackhouse, G. and Briggs, A. and O'Brien, B.J. (2002) A note on the estimation of confidence intervals for cost-effectiveness when costs and effects are censored. *Medical Decision Making* 22(2):pp. 173-177.

A Note on the Estimation of Confidence Intervals for Cost-Effectiveness When Costs and Effects Are Censored*

Gordon Blackhouse, MBA, Andrew H. Briggs, DPhil, Bernie J. O'Brien, PhD

Background. *The relation between methodological advances in estimation of confidence intervals (CIs) for incremental cost-effectiveness ratios (ICER) and estimation of cost effectiveness in the presence of censoring has not been explored. The authors address the joint problem of estimating ICER precision in the presence of censoring.* ***Methods***. *Using patient-level data (n = 168) on cost and survival from a published placebo-controlled trial, the authors compared 2 methods of measuring uncertainty with censored data: 1) Bootstrap with censor adjustment (BCA); 2) Fieller's method with censor adjustment (FCA). The authors estimate the FCA over all possible values for the correlation (ρ) between costs and effects (range= –1 to +1) and also examine the use of the correlation between cases without censoring adjustment (i.e., simple time-on-study) for costs and effects as an approximation for ρ.* ***Results***. *Using time-on-study, which considers all censored observations as responders (deaths), yields 0.64 life-years gained at an additional cost of 87.9 for a cost per life-year of 137 (95% CI by bootstrap –5.9 to 392). Censoring adjustment corrects for the bias in the time-on-study approach and reduces the cost per life-year estimate to 132 (=72/0.54). Confidence intervals with censor adjustment were approximately 40% wider than the base-case without adjustment. Using the Fieller method with an approximation of ρ based on the uncensored cost and effect correlation provides a 95% CI of (–48 to 529), which is very close to the BCA interval of (–52 to 504).* ***Conclusions***. *Adjustment for censoring is necessary in cost-effectiveness studies to obtain unbiased estimates of ICER with appropriate uncertainty limits. In this study, BCA and FCA methods, the latter with approximated covariance, are simple to compute and give similar confidence intervals.* ***Key words:*** *confidence intervals; censored data; cost effectiveness analysis; Fieller's theorem; bootstrap.*

The growing number of prospective economic evaluations alongside clinical trials has raised several methodological issues on the statistical analysis of cost and cost-effectiveness data. Two areas of recent innovation are methods for estimating confidence limits for incremental cost-effectiveness[1–4] and methods for the analysis of censored cost data where follow-up is incomplete.[5–7] Confidence intervals for cost-effectiveness can be estimated nonparametrically using the resampling technique of bootstrapping[3,8] or parametrically using Fieller's theorem[2,9] where cost and effect differences are assumed to follow a bivariate normal distribution. Censoring due to incomplete follow-up is a problem for unbiased estimation of mean effects and costs. Building on techniques for analyzing censored effect data, Lin et al.[6] and Etzioni et al.[5] developed the Kaplan-Meier Sample Average (KMSA) method. This method has been shown to provide a consistent estimator of cost in the presence of censoring.

With the exception of earlier work by Gardiner et al.[10] who examined these problems in relation to deterministic (i.e., nonrandom) cost functions, the methodological advances in estimation of confidence intervals and censored costs have been made in apparent isolation from one another: the work on confidence inter-

---

Received 12 February 2001 from the Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada (GB, BJO), the Centre for Evaluation of Medicines, St Joseph's Hospital, Hamilton (GB, AHB, BJO), and Health Economics Research Centre, Institute of Health Sciences, University of Oxford, UK (AHB). Revision accepted for publication 8 November 2001. This article was prepared while Dr. Briggs was visiting the Centre for Evaluation of Medicines (CEM) at McMaster University, funded through a joint U.K. Medical Research Council/Southeast Region Training Fellowship and the Centre for Evaluation of Medicines.

Address correspondence and reprint requests to Gordon Blackhouse, Centre for Evaluation of Medicines, St. Joseph's Hospital, 50 Charlton Avenue East, H-305, Hamilton, Ontario, Canada L8N 4A6; telephone: (905) 522-1155, ext. 5214; e-mail: blackhou@mcmaster.ca.

*Presented at the Annual Meeting of the Society for Medical Decision Making, October 2000, Cincinnati, Ohio.

vals assuming that uncensored data are available[1–3,9] and the work on censoring adjustments focusing on costs and effects separately.[6,7,11,12] In this article, we use data on costs and effects from a clinical trial to explore 2 approaches to the joint problem of estimating confidence intervals with censored cost and effect data:

Censor-Adjusted Bootstrap Interval: achieved by bootstrapping the KMSA estimator for costs and effects simultaneously to determine their joint density;

Censor-Adjusted Fieller's Interval: achieved by applying the KMSA estimator for costs and effects separately and using these estimates in Fieller's theorem with an approximated covariance between costs and effects.

METHODS

Cost-Effectiveness Trial

Data were available from a randomized trial of an antiviral therapy where patient-specific data on resource consumption, length of follow-up, and death status at end of study were obtained for 168 patients, with $n = 67$ assigned to control and $n = 81$ assigned to treatment. Costs were calculated by weighting resource use by unit costs (although the currency of costing is not specified). These trial data are right-censored because time-to-death is the dependent variable and not all individuals have "responded" (died) in the observation period, so their final cost and survival is unknown. In the control arm, 22% of observations were censored, whereas 29% of observations in the treatment arm were censored. To illustrate the importance of censoring adjustment, we 1st estimated cost-effectiveness with no adjustment for censoring and assumed that all patients "respond" (die) at the end of their observed time-on-study. This method yields biased (downward) estimates of mean costs and effects.

Life Expectancy

Kaplan-Meier (K-M) survival curves[13] for time-to-death were estimated and, following convention,[14] life expectancy (mean survival) for treatment and control groups was calculated as the area under the survival curve. The time horizon for the survival curves for both treatment and control groups was set to the length of follow-up of the last uncensored observation (1307 days). Standard errors for the K-M estimators for mean survival times were used to estimate a 95% confidence interval for the difference in life expectancy.

Kaplan-Meier Sample Average (KMSA) Censor Adjustment for Costs

Following the method described by Etzioni et al.[5] to calculate the KMSA estimator for costs, the trial period is divided into k time periods. The mean costs of all uncensored patients during time intervals $t = 1$ to k is multiplied by the Kaplan-Meier estimate of the proportion alive at the beginning of the interval. Summing across these weighted costs for each interval gives the KMSA estimator of costs adjusted for censoring. The KMSA is a consistent estimator of mean cost when censoring occurs only at the beginning of the interval and is nearly consistent for small time intervals for any other censoring pattern.[6] We use 90-day intervals in the estimation of the KMSA in our analysis. Lin and colleagues have provided an estimate of the standard error for these costs so that a confidence interval for the cost-difference is straightforward to calculate using standard methods.[6]

Censor-Adjusted Bootstrap Interval

The key difference between use of bootstrapping for uncensored and censored cost-effectiveness data is that the latter circumstance requires additional information on whether the subject is censored at observed follow-up time t, these data being part of the resampling frame. The procedure also makes greater computational demands, working with censoring adjusted estimators (e.g., K-M for survival and KMSA for costs) rather than with simple sample means.

We followed the following procedure:

A sample was drawn (with replacement) of 67 patients from the control group, with data elements per patient of survival time, cost and censoring status (yes/no). This formed a single bootstrap resample of the control group. A similar sample ($n = 81$) was drawn from the treatment group to form a single bootstrap resample of the treatment group.

1. A Kaplan-Meier estimate of censor-adjusted mean survival (life expectancy) and a KMSA estimate for censor-adjusted costs was calculated from the bootstrap resamples of the control and treatment groups.
2. The censor-adjusted cost and effect replicates from each treatment and control bootstrap resample are then used to calculate an incremental cost-effectiveness ratio (ICER)—this being one replicate of the ICER.

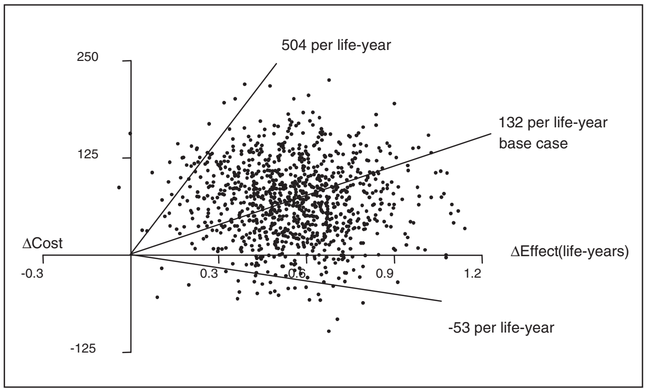

Figure 1. Cost and effect pairs from 1000 bootstrap samples showing base case cost-effectiveness and 95% confidence intervals (CIs).

Repeating these 2 steps 1000 times generated a vector of bootstrap replicates of cost-effectiveness results. The 1000 cost and effect estimates used to calculate the ICER replicates form the nonparametric estimate of the joint density of costs and effects.

Censor-Adjusted Fieller's Interval

The method for estimating a confidence interval for an ICER using Fieller's theorem has been presented in several publications, but always in the context of uncensored data.[1,2,9] Based on the literature reviewed above, it is clearly possible to derive the relevant censor-adjusted means and variances for costs and effects (KMSA and K-M Life Expectancy estimates for treatment and control groups) for input to the Fieller confidence interval formula.

In the absence of an exact solution for the covariance term with censoring, which is required for the Fieller method, we present a pragmatic way forward. We use K-M and KMSA estimators for mean survival and costs, respectively, for treatment and control groups, and with relevant variance terms for input into the Fieller confidence interval formula. Covariance of incremental costs and effects can be expressed as a function of variance of incremental costs, variance of incremental effects, and correlation between incremental costs and effects: $\mathrm{cov}(\Delta E, \Delta C) = \rho * \sqrt{\mathrm{var}(\Delta E) * \mathrm{var}(\Delta C)}$. Because we have inputs necessary to estimate the variance of incremental costs and effects, an indirect method of estimating the covariance of incremental costs and effects is to assume a specific correlation (ρ) between costs and effects.

For the correlation between censored costs and effects (ρ), we examined 2 options: 1) an assumed effect-cost correlation ranging from –1 to +1, including 0 for cost and effect independence; 2) an approximated correlation based on the Pearson correlation between the cost and effect data without censoring adjustment (i.e., time-on-study with all patients assumed to "respond").

RESULTS

Using observed time-on-study and assuming all censored observations are responders (deaths) yields survival of 1.34 years with control and 1.98 years with treatment; difference of 0.64 years in favor of treatment (95% CI 0.30 to 0.98). The K-M estimates (without censor adjustment) for mean survival were 1.53 years with control and 2.07 years with treatment, a difference in favor of treatment of 0.54 years (95% CI 0.16 to 0.92). Mean cost per patient without censoring adjustments was 363.9 with control and 451.8 with treatment; a cost increase with treatment of 87.9 (95% CI –5.0 to 180.8). Adjusting costs for censoring, the KSMA estimates are 395.8 with control and 467.8 with treatment, a cost increase with treatment of 72 (95% CI –23.4 to 167.4). Overall, the ICER with no censor adjustments is 139 and with censor adjustment, this falls to 132.

The 1000 bootstrap replicates of censor-adjusted cost and effect differences are plotted in Figure 1, and the bootstrap replicates of the associated ICERs are represented by the slopes of the lines joining these points to the origin. Confidence intervals for the ICER are estimated from the 2.5th and 97.5th percentiles of a rank ordering of the ICER replicates, with negative ICERs in the northwest quadrant of the plane appearing at the top of the rank order. The estimated percentile confidence interval limits are –52 to 504, the negative lower limit indicating treatment could dominate control (less costly and more effective).

The ICER confidence intervals for the bootstrap and Fieller methods with censoring (and for bootstrap without censoring) are presented in Table 1. Comparing the bootstrap method with and without censoring adjustment indicates little impact on the point estimate of the ICER, but censoring adjustment widens the confidence interval by 40% (from 398 to 556). The Fieller confidence interval depends on the assumed correlation between costs and effects; with perfect negative correlation, the length of the interval is 1017 (from limits of –25.5 to 991.5), and with perfect positive correlation, the length of the interval is 596 (from limits of –45.9 to 550.4).

Table 1. Estimates of 95% Confidence Intervals (CI) for Incremental Cost-Effectiveness with and without Adjustments for Censoring (antiviral treatment data from Fenn et al.[11])

Method	Adjustment for Censoring?	Incremental Cost-Effectiveness Ratio (cost per life-year gained) Base Case	Lower 95% CI	Upper 95% CI	Interval
Bootstrap	No	137.2	–5.9	392.0	397.9
Bootstrap	Yes	132.2	–52.5	503.7	556.2
Fieller ($\rho = +1$)[a]	Yes	132.2	–138.8	182.0	320.8
Fieller ($\rho = 0$)[a]	Yes	132.2	–45.9	550.4	596.3
Fieller ($\rho = -1$)[a]	Yes	132.2	–25.5	991.3	1017.8
Fieller ($\rho = 0.05$)[b]	Yes	132.2	–47.7	529.2	576.9

Note: ρ = correlation between differences in cost and effects, which is assumed for 3 Fieller methods ([a]) and approximated by the correlation between uncensored costs and effects for the final Fieller method ([b]).

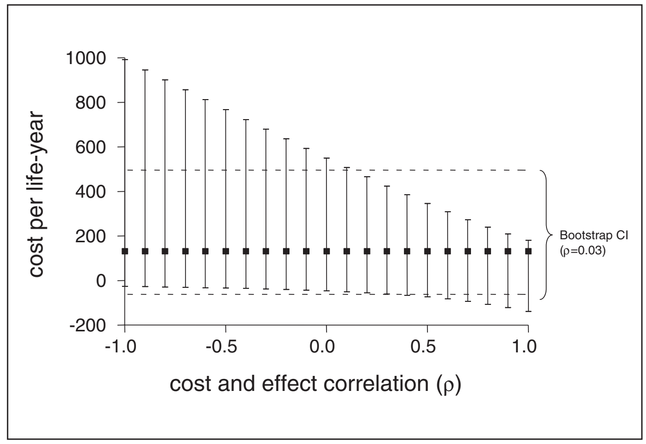

Figure 2. Incremental cost-effectiveness (■) and 95% confidence interval (CI) by Fieller's method with alternative values for effect-cost correlation (ρ).

The relation between alternative assumptions on cost and effect correlation and the ICER confidence interval is plotted in Figure 2. The size and symmetry of the confidence interval varies over the range of correlation from –1 (large interval, positive skew) to +1 (small interval, negative skew). In both Figure 2 and Table 1 is shown the Fieller interval with approximated covariance based on uncensored cost-effect correlation ($\rho = 0.05$); this is very similar to the bootstrap interval with censor adjustment, shown by the dashed lines in Figure 2. The cost-effect correlation calculated from the censor adjusted bootstrap replicates is 0.03.

DISCUSSION

This is the first study to present bootstrap and Fieller methods for cost-effectiveness confidence intervals with censored cost and effect data. Our data comparing the uncensored and censored bootstrap data reveal that adjustments for censoring have important implications for both the bias and precision of cost-effectiveness estimates. Although the difference in point estimate (bias) was small, the size of the confidence interval is substantially increased when necessary adjustments are made for censoring.

The choice between the censor-adjusted bootstrap and Fieller method for confidence interval estimation rests on a number of factors. The computer-intensive challenge of bootstrapping is no longer a major constraint in the era of high-speed computing. The bootstrapping algorithm involves the re-estimation of multiple Kaplan-Meier curves and KMSA estimators and is not trivial to set up, but once achieved it offers a robust method, which preserves the correlation structures within the data. Although the size and symmetry of the Fieller's confidence interval are clearly a function of the assumed covariance (Figure 2), the similarity between the Fieller method when employing the unadjusted correlation and the bootstrap results suggests this may be a pragmatic way forward if analysts want to exploit the relative simplicity of Fieller's approach.

REFERENCES

1. Briggs AH, Fenn P. Confidence intervals or surfaces? Uncertainty on the cost-effectiveness plane. Health Econ. 1998;7(8):723–40.

2. Willan A, O'Brien BJ. Confidence intervals for cost-effectiveness ratios in clinical trials: a new method using Fieller's theorem. Health Econ. 1996;5:297–305.

3. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Econ. 1997;6:327–40.

4. O'Brien BJ, Drummond MF, Labelle RJ, Willan A. In search of power and significance: issues in the design and analysis of stochastic cost-effectiveness studies in health care. Med Care. 1994;32(2): 150–63.

5. Etzioni R, Urban N, Baker M. Estimating the costs attributable to a disease with application to ovarian cancer. J Clin Epidemiol. 1996;32(2):95–103.

6. Lin DY, Feuer EJ, Etzioni R, Wax Y. Estimating medical costs from incomplete follow-up data. Biometrics. 1997;53(2):419–34.

7. Fenn P, McGuire A, Phillips V, Backhouse M, Jones D. The analysis of censored treatment cost data in economic evaluation. Med Care. 1995;33(851):863.

8. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York: Chapman and Hall, 1997.

9. Chaudhary MA, Stearns SC. Estimating confidence intervals for cost-effectiveness ratios: an example from a randomized trial. Stat Med. 1996;15:1447–58.

10. Gardiner J, Hogan A, Holmes-Rovner M, Rovner D, Griffith L, Kuppersmith J. Confidence intervals for cost-effectiveness ratios. Med Decis Making. 1995;15:254–63.

11. Fenn P, McGuire A, Backhouse M, Jones D. Modelling programme costs in economic evaluation. J Health Econ. 1996;15:115–25.

12. Etzioni R, Feuer EJ, Sullivan S, Lin DY, Hu C, Ramsey S. On the use of survival analysis to estimate medical care costs. Health Econ. 1999;18:365–80.

13. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958;(53):457–81.

14. Wright JC, Weinstein MC. Gains in life expectancy from medical interventions—standardizing data on outcomes. N Engl J Med. 1998;339(6):380–6.