UW Biostatistics Working Paper Series

3-30-2005

Nonparametric Statistical Methods for the Cost-Effectiveness Analysis

Phillip Dinh
University of Washington, ispvd@u.washington.edu

Xiao-Hua Zhou
University of Washington, azhou@u.washington.edu

1. Introduction

Cost-effectiveness analyses are commonly used techniques in health services research. In a cost-effectiveness analysis, two groups of subjects are often compared; e.g., new intervention against existing intervention or treatment A versus treatment B. As its name suggests, a cost-effectiveness analysis considers both cost and effectiveness of the interventions. The two measures commonly used in a cost-effectiveness analysis are the incremental cost-effectiveness ratio (ICER) and the net health benefit (NHB). Each measure has its own advantages and disadvantages as we summarize in the list below. In this paper, we do not recommend one approach over the other. For this reason, we discuss statistical methods for both measures.

Formally, let the cost and effectiveness for subject j who receives treatment i be C_{ij} and E_{ij}, respectively. Let us further assume that (C_{ij}, E_{ij}) comes from a bivariate distribution with mean (μ_{C_i}, μ_{E_i}), variance $(\sigma^2_{C_i}, \sigma^2_{E_i})$ and covariance $\sigma_{E_iC_i}$. That is, for i = 1, 2,

$$\begin{pmatrix} C_{i1} \\ E_{i1} \end{pmatrix}, \begin{pmatrix} C_{i2} \\ E_{i2} \end{pmatrix}, \ldots, \begin{pmatrix} C_{in_i} \\ E_{in_i} \end{pmatrix} \sim_{i.i.d} \left(\begin{pmatrix} \mu_{C_i} \\ \mu_{E_i} \end{pmatrix}, \begin{pmatrix} \sigma^2_{C_i} & \sigma_{C_iE_i} \\ \sigma_{C_iE_i} & \sigma^2_{E_i} \end{pmatrix} \right). \quad (1)$$

If λ is the willingness-to-pay per unit of the effectiveness (Stinnett and Mullahy, 1998), the ICER and the NHB are defined by

$$ICER = \frac{\mu_{C_1} - \mu_{C_2}}{\mu_{E_1} - \mu_{E_2}}, \quad NHB = (\mu_{E_1} - \mu_{E_2}) - \frac{1}{\lambda}(\mu_{C_1} - \mu_{C_2}),$$

respectively.

Using sample data, these two quantities can be estimated as:

$$\widehat{ICER} = \frac{\bar{C}_1 - \bar{C}_2}{\bar{E}_1 - \bar{E}_2}, \quad \widehat{NHB} = (\bar{E}_1 - \bar{E}_2) - \frac{1}{\lambda}(\bar{C}_1 - \bar{C}_2),$$

respectively, where $\bar{C}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} C_{ij}$ and $\bar{E}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} E_{ij}$ (i = 1, 2).

- **Arguments for the use of the ICER include:**

- The ICER has a natural interpretation as a price-per-unit of product (e.g., $15,000 per quality-adjusted life year saved).

- Analysis on the ICER is independent of the cut-off (willingness-to-pay) λ value. This cut-off value will vary from context to context depending on preferences and the available budgets (Briggs and Fenn, 1998).

- The ICER has an economic foundation based on utility theory. Maximizing expected utility leads to finding optimal incremental cost-effectiveness ratio (Garber and Phelps, 1997).

- **Arguments against the use of the ICER include:**

- Interpretation of a negative ICER is problematic. First, the magnitude of negative ICERs is meaningless. An intervention that is preferred may have smaller (negative) ICER (Briggs and Fenn, 1998). Second, both quadrants II and IV on the cost-effectiveness plane (Figure 1) generate negative ICERs; however, decision-makings are exactly opposite (Stinnett and Mullahy, 1998).

- Confidence intervals for the ICER can include an undefined value. Figure 1 shows a cost-effectiveness plane with the confidence interval represented by the two broken lines. This interval includes the vertical axis where $\mu_{E_1} - \mu_{E_2} = 0$ and an undefined value for the ICER (Willan, 2003).

[Figure 1 here]

- Statistically, $\widehat{ICER}$ is not a sufficient statistic and $\widehat{ICER}$ is biased for ICER (Zethraeus et al., 2003).

- **Arguments for the use of the NHB include:**

- The NHB is properly ordered. Treatment with the largest NHB is the most cost-effective (Willan, 2003).

- Interpretation for the NHB is not ambiguous like the negative ICER.

- Since the NHB is linear in both costs and effects, $\widehat{NHB}$ is unbiased for NHB. Also, asymptotic results may apply at smaller sample sizes.

- The NHB can be extended easily to more than two treatments.

• **Arguments against the use of the NHB include:**

- The NHB incorporates λ into the analysis. This λ may vary from setting to setting and the analyst's λ may be different from the policy maker's λ.

- The NHB does not have a natural interpretation as price per unit of product like the ICER.

Nonparametric statistical methods have been developed for inferences on the two measures. However, due to the skew nature of cost data, existing methods have not adequately addressed the problems. In this paper, we are interested in studying the theoretical performance of normal-based intervals and constructing new confidence intervals for the ICER and the NHB.

1.1 *Existing Methods*

Confidence intervals for the ICER and the NHB have been constructed using several existing techniques. For the ICER, intervals based on the asymptotic theory, the Fieller's (Fieller, 1954), and several versions of the bootstrap have been adopted for inferences on the ICER. For reviews and evaluations of these techniques, see Briggs et al. (1999) and Fan and Zhou (2005).

For the NHB, similar techniques have been proposed. They include the normal interval based on large sample theory and the bootstrap intervals (Stinnett and Mullahy, 1998; Willan, 2001).

1.2 *Limitations of Existing Methods*

For the ICER, normal theory intervals and bootstrap intervals which assume normal sampling distribution have poor coverages (Briggs et al., 1999; Fan and Zhou, 2005). In term of coverage accuracy, Briggs et al. (1999) found the Fieller's interval was best while Fan and Zhou (2005) recommended the bootstrap-t interval. However, Fieller's interval requires finding roots of a quadratic equation and these can be unreal. In addition, if this quadratic equation has one root, the confidence interval will be half-open. In our simulation study presented later, we also found the Fieller's interval may not give best coverage. The bootstrap-t interval gives good coverage but is often too wide. When having such wide intervals, one has to consider the trade off between the coverage accuracy and the precision of the estimate.

For the NHB, inferences based on large sample theory may not be appropriate for highly skewed data. Our simulation study indicates that coverage of such normal theory intervals can differ significantly from the nominal value.

1.3 *Proposed Approach*

Beside the above methods, another approach is to modify the test statistics to reduce the effect of skewness. The method is based on the Edgeworth expansion (Hall, 1992a). We will follow this approach in this paper.

In Sections 2 and 3, we will develop the Edgeworth expansions for the studentized t-statistics for the ICER and the NHB, respectively. We will use the expansions to study the theoretical performance of existing normal theory intervals and to derive new transformational intervals to improve coverage accuracy. In Section 4, we will demonstrate the method via a simulation study. In Section 5, we will apply our method to a real cost data set. In

Section 6, we will summarize the methods and provide our recommendation.

2. Edgeworth expansion for the incremental cost-effectiveness ratio (ICER)

Cost-effectiveness analyses are often hindered by skewed cost data. In this section, we derive the Edgeworth expansion for the studentized t-statistic for the ICER. The expansion will then be used to guide inferences. In addition, we use the expansion to derive new transformational intervals for the ICER.

Assume our data is given as in equation (1). The ICER is defined by $\theta_1 \equiv ICER = \frac{\mu_{C_1}-\mu_{C_2}}{\mu_{E_1}-\mu_{E_2}}$ and is estimated by $\hat{\theta}_1 \equiv \widehat{ICER} = \frac{\bar{C}_1-\bar{C}_2}{\bar{E}_1-\bar{E}_2}$. The asymptotic variance of the $\widehat{ICER}$ can be estimated by

$$\hat{\sigma}_1^2 = \frac{\left[\frac{S^2_{C_1}}{n_1}+\frac{S^2_{C_2}}{n_2}\right]}{(\bar{E}_1-\bar{E}_2)^2} + \frac{(\bar{C}_1-\bar{C}_2)^2\left[\frac{S^2_{E_1}}{n_1}+\frac{S^2_{E_2}}{n_2}\right]}{(\bar{E}_1-\bar{E}_2)^4} - \frac{2(\bar{C}_1-\bar{C}_2)\left[\frac{S_{E_1C_1}}{n_1}+\frac{S_{E_2C_2}}{n_2}\right]}{(\bar{E}_1-\bar{E}_2)^3} \quad (2)$$

where, for i = 1, 2, $\bar{C}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} C_{ij}$, $\bar{E}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} E_{ij}$, $S^2_{C_i} = \frac{1}{n_i}\sum_{j=1}^{n_i}(C_{ij} - \bar{C}_i)^2$, $S^2_{E_i} = \frac{1}{n_i}\sum_{j=1}^{n_i}(E_{ij}-\bar{E}_i)^2$, and $S_{C_iE_i} = \frac{1}{n_i}\sum_{j=1}^{n_i}(C_{ij}-\bar{C}_i)(E_{ij}-\bar{E}_i)$.

Theorem 1 Let $\nu_N = n_1/(n_1+n_2) = n_1/N$. Assume $\nu_N = \nu + O(N^{-r})$ for some $r \geq 0$. Let $T_1 = \frac{\widehat{ICER}-ICER}{\hat{\sigma}_1}$. Under regularity conditions (Hall, 1992a), the distribution of T_1 has the following expansion,

$$P(T_1 \leq x) = \Phi(x) + N^{-1/2}q_1(x)\phi(x) + O(N^{-min(1,r+1/2)}) \quad (3)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and distribution functions, respectively, of the standard normal variable, and, with $E(\cdot)$ is the expectation function,

$q_1(x) = \frac{A_1}{3} + \frac{2A_1+3A_2}{3}x^2$,

$A_1 = \frac{1}{2\Delta_1^6 Q_1^{3/2}}\left[\frac{P_1}{\nu^2} - \frac{P_2}{(1-\nu)^2}\right]$, $A_2 = \frac{(\Delta_2\Gamma_1 - \Delta_1\Gamma_{12})}{\Delta_1^3 Q_1^{1/2}}$,

$P_1 = E[\Delta_1(C_{1j}-\mu_{C_1}) - \Delta_2(E_{1j}-\mu_{E_1})]^3$,

$P_2 = E[\Delta_1(C_{2j}-\mu_{C_2}) - \Delta_2(E_{2j}-\mu_{E_2})]^3$,

$Q_1 = \frac{\Gamma_2}{\Delta_1^2} + \frac{\Gamma_1 \Delta_2^2}{\Delta_1^4} - 2\frac{\Delta_2 \Gamma_{12}}{\Delta_1^3}, \;\; \Delta_1 = \mu_{E_1} - \mu_{E_2}, \;\; \Delta_2 = \mu_{C_1} - \mu_{C_2},$
$\Gamma_1 = \frac{\sigma^2_{E_1}}{\nu} + \frac{\sigma^2_{E_2}}{1-\nu}, \quad \Gamma_2 = \frac{\sigma^2_{C_1}}{\nu} + \frac{\sigma^2_{C_2}}{1-\nu}, \quad \Gamma_{12} = \frac{\sigma_{E_1 C_1}}{\nu} + \frac{\sigma_{E_2 C_2}}{1-\nu}.$

For a proof, see Appendix A1.

From the expansion (3), we see that $q_1(x)/\sqrt{N}$, in absoluate value, plays an important role in the normal approximation of T_1. When $|q_1(x)|/\sqrt{N}$ is small, T_1 can be approximated by a normal distribution accurately. On the contrary, when $|q_1(x)|/\sqrt{N}$ is large, the second term in equation (3) is not ignorable, thus the normal approximation won't be as accurate. The term $q_1(x)$ can be large when either A_1 or A_2, or both are large. The quantity A_1 relates directly to the skewness of the cost data, thus can be large when the cost data is highly skewed. The term A_2 relates to the variance and covariance of the data, which can be large when the data is highly variable.

The expansion can be used to construct transformational confidence intervals for the ICER. The intervals will correct for the term $q_1(x)$ in the expansion (3). Following the similar ideas as in Hall (1992b) and Zhou and Dinh (2005), with

$$g_1^{-1}(x) = (\hat{\gamma})^{-1}[1 + 3\hat{\gamma}(x + \hat{\gamma}/N)]^{1/3} - (\hat{\gamma})^{-1},$$
$$g_2^{-1}(x) = (2N^{-1/2}\hat{\gamma})^{-1} log[2N^{-1/2}\hat{\gamma}(x + N^{-1}\hat{\gamma}) + 1],$$
$$g_3^{-1}(x) = [1 + 3(x + \hat{\gamma}/N)]^{1/3} - 1,$$

we proposed three transformational intervals,

$$\hat{\theta}_1 - \hat{\sigma}_1 \left[\frac{\hat{c}_1}{\sqrt{N}} + \sqrt{N} g_i^{-1}\left(\frac{z_{1-\alpha/2}}{\sqrt{N}}\right)\right] \leq \theta_1 \leq \hat{\theta}_1 - \hat{\sigma}_1 \left[\frac{\hat{c}_1}{\sqrt{N}} + \sqrt{N} g_i^{-1}\left(\frac{z_{\alpha/2}}{\sqrt{N}}\right)\right] \quad (4)$$

where, i=1, 2, 3, $z_\alpha = \Phi(\alpha), \;\; \hat{\gamma} = \frac{2\hat{A}_1 + 3\hat{A}_2}{3}, \;\; \hat{c}_1 = -\hat{A}_1 - \hat{A}_2,$
$\hat{A}_1 = \frac{1}{2\hat{\Delta}_1^6 \hat{Q}_1^{3/2}} \left[\frac{\hat{P}_1}{\nu_N^2} - \frac{\hat{P}_2}{(1-\nu_N)^2}\right], \;\; \hat{A}_2 = \frac{(\hat{\Delta}_2 \hat{\Gamma}_1 - \hat{\Delta}_1 \hat{\Gamma}_{12})}{\hat{\Delta}_1^3 \hat{Q}_1^{1/2}},$
$\hat{P}_1 = \frac{1}{n_1} \sum_{j=1}^{n_1} [\hat{\Delta}_1 (C_{1j} - \bar{C}_1) - \hat{\Delta}_2 (E_{1j} - \bar{E}_1)]^3,$

$\hat{P}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2}[\hat{\Delta}_1(C_{2j}-\bar{C}_2)-\hat{\Delta}_2(E_{2j}-\bar{E}_2)]^3$,
$\hat{Q}_1 = \frac{\hat{\Gamma}_2}{\hat{\Delta}_1^2}+\frac{\hat{\Gamma}_1\hat{\Delta}_2^2}{\hat{\Delta}_1^4}-2\frac{\hat{\Delta}_2\hat{\Gamma}_{12}}{\hat{\Delta}_1^3}$, $\hat{\Delta}_1 = \bar{E}_1-\bar{E}_2$, $\hat{\Delta}_2 = \bar{C}_1-\bar{C}_2$,
$\hat{\Gamma}_1 = \frac{S_{E_1}^2}{\nu_N}+\frac{S_{E_2}^2}{1-\nu_N}$, $\hat{\Gamma}_2 = \frac{S_{C_1}^2}{\nu_N}+\frac{S_{C_2}^2}{1-\nu_N}$, $\hat{\Gamma}_{12} = \frac{S_{E_1C_1}}{\nu_N}+\frac{S_{E_2C_2}}{1-\nu_N}$.

For the derivation, see Appendix A2.

3. Edgeworth expansion for the net health benefit (NHB)

In this section, we present the Edgeworth expansion for the studentized t-statistic for the NHB. The expansion provides a way to correct for the skewness in cost data and to derive new confidence intervals for the NHB.

Let the data be given as in equation (1), the NHB is defined by $\theta_2 \equiv NHB = (\mu_{E_1}-\mu_{E_2})-\frac{1}{\lambda}(\mu_{C_1}-\mu_{C_2})$ and can be estimated as $\hat{\theta}_2 \equiv \widehat{NHB} = (\bar{E}_1-\bar{E}_2)-\frac{1}{\lambda}(\bar{C}_1-\bar{C}_2)$. The asymptotic variance of $\hat{\theta}_2$ can be estimated by,

$$\hat{\sigma}_2^2 = \frac{S_{E_1}^2}{n_1}+\frac{S_{E_2}^2}{n_2}+\frac{S_{C_1}^2}{\lambda^2 n_1}+\frac{S_{C_2}^2}{\lambda^2 n_2}-\frac{2S_{C_1E_1}}{\lambda n_1}-\frac{2S_{C_2E_2}}{\lambda n_2} \quad (5)$$

Theorem 2 Let $\nu_N = n_1/(n_1+n_2) = n_1/N$. Assume $\nu_N = \nu + O(N^{-r})$ for some $r \geq 0$. Let $T_2 = \frac{\widehat{NHB}-NHB}{\hat{\sigma}_2}$. Under regularity conditions (Hall, 1992a), the distribution of T_2 has the following expansion,

$$P(T_2 \leq x) = \Phi(x) + N^{-1/2}q_2(x)\phi(x) + O(N^{-min(1,r+1/2)}) \quad (6)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and distribution functions, respectively, of the standard normal variable,

$q_2(x) = \frac{A}{6}(2x^2+1)$,
$Q_2 = \frac{\sigma_{E_1}^2}{\nu}+\frac{\sigma_{E_2}^2}{(1-\nu)}+\frac{\sigma_{C_1}^2}{\lambda^2\nu}+\frac{\sigma_{C_2}^2}{\lambda^2(1-\nu)}-\frac{2\sigma_{C_1E_1}}{\lambda\nu}-\frac{2\sigma_{C_2E_2}}{\lambda(1-\nu)}$,
$B_1 = E\left(\frac{C_{1j}-\mu_{C_1}}{\lambda}-(E_{1j}-\mu_{E_1})\right)^3$, $B_2 = E\left(\frac{C_{2j}-\mu_{C_2}}{\lambda}-(E_{2j}-\mu_{E_2})\right)^3$,

where $E(\cdot)$ is the expectation function, and

$$A = \frac{1}{Q_2^{3/2}}\left[\frac{B_1}{\nu^2}-\frac{B_2}{(1-\nu)^2}\right]. \quad (7)$$

For a proof, see appendix B.

Similar to the ICER, we see that the normal approximation to T_2 is accurate when $|q_2(x)|/\sqrt{N}$ is small. When $|q_2(x)|/\sqrt{N}$ is large, the second term in the expansion (6) can't be ignored and the normal approximation won't be as accurate. The quantity $q_2(x)$ relates directly to the skewness of the cost data, thus can be large when the cost data is highly skewed.

Similar to the ICER, we can derive the three transformational confidence intervals for the NHB. The intervals will correct for the term $q_2(x)$ in the expansion (6) above. With,

$$g_1^{-1}(x) = \left(\tfrac{1}{3}\hat{A}\right)^{-1}\left[1 + 3\tfrac{1}{3}\hat{A}(x - N^{-1}\tfrac{1}{6}\hat{A})\right]^{1/3} - \left(\tfrac{1}{3}\hat{A}\right)^{-1},$$
$$g_2^{-1}(x) = \left(2\tfrac{1}{3}N^{-1/2}\hat{A}\right)^{-1} log\left[2\tfrac{1}{3}N^{-1/2}\hat{A}(x - N^{-1}\tfrac{1}{6}\hat{A}) + 1\right],$$
$$g_3^{-1}(x) = \left[1 + 3(x - N^{-1}\tfrac{1}{6}\hat{A})\right]^{1/3} - 1,$$

the $(1-\alpha)100\%$ transformational confidence interval for the NHB is given by

$$\widehat{NHB} - \sqrt{N}g_i^{-1}\left(\frac{z_{1-\alpha/2}}{\sqrt{N}}\right)\hat{\sigma}_2 \le NHB \le \widehat{NHB} - \sqrt{N}g_i^{-1}\left(\frac{z_{\alpha/2}}{\sqrt{N}}\right)\hat{\sigma}_2 \quad (8)$$

where, i = 1, 2, 3, $z_\alpha = \Phi(\alpha)$, $\hat{A} = \frac{1}{\hat{Q}_2^{3/2}}\left[\frac{\hat{B}_1}{\nu_N^2} - \frac{\hat{B}_2}{(1-\nu_N)^2}\right]$,
$\hat{Q}_2 = \frac{S_{E_1}^2}{\nu_N} + \frac{S_{E_2}^2}{(1-\nu_N)} + \frac{S_{C_1}^2}{\lambda^2\nu_N} + \frac{S_{C_2}^2}{\lambda^2(1-\nu_N)} - \frac{2S_{C_1E_1}}{\lambda\nu_N} - \frac{2S_{C_2E_2}}{\lambda(1-\nu_N)}$,
$\hat{B}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1}\left(\frac{C_{1j}-\bar{C}_1}{\lambda} - (E_{1j}-\bar{E}_1)\right)^3$, $\hat{B}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2}\left(\frac{C_{2j}-\bar{C}_2}{\lambda} - (E_{2j}-\bar{E}_2)\right)^3$.

4. Simulation study

4.1 *ICER Simulation*

In this section we report a simulation study to evaluate the method presented in Sections 2. In particular, we compared our new intervals given in equation (4) against the Taylor's interval (based on normal theory), the

Fieller's interval (recommended by Briggs et al. (1999)), the bootstrap percentile-t interval (recommended by Fan and Zhou (2005)), and the Hinkley's interval. Details of these methods are given in Appendix C.

We generated data from three bivariate distributions: bivariate normal, bivariate lognormal, and bivariate mixture (costs are lognormal and effectiveness data are normal). We chose different correlation structures for costs and effects and sample sizes as presented in Tables 1 and 2. The data were generated as following:

Bivariate normal:

$$\begin{pmatrix} C_{1j} \\ E_{1j} \end{pmatrix} \sim_{i.i.d} N_2\left(\begin{pmatrix} \mu_1 = 40,000 \\ \mu_2 = 60 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 = 200,000 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 = 5 \end{pmatrix}\right),$$
$$\begin{pmatrix} C_{2j} \\ E_{2j} \end{pmatrix} \sim_{i.i.d} N_2\left(\begin{pmatrix} \mu_3 = 30,000 \\ \mu_4 = 50 \end{pmatrix}, \begin{pmatrix} \sigma_3^2 = 100,000 & \sigma_{34} \\ \sigma_{34} & \sigma_4^2 = 10 \end{pmatrix}\right),$$

Bivariate mixture:

$$\begin{pmatrix} C_{1j} \\ E_{1j} \end{pmatrix} = \begin{pmatrix} e^{Y_{C1j}} \\ Y_{E1j} \end{pmatrix}, \begin{pmatrix} Y_{C1j} \\ Y_{E1j} \end{pmatrix} \sim_{i.i.d} N_2\left(\begin{pmatrix} \mu_1 = 8 \\ \mu_2 = 4 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 = 2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 = .5 \end{pmatrix}\right),$$
$$\begin{pmatrix} C_{2j} \\ E_{2j} \end{pmatrix} = \begin{pmatrix} e^{Y_{C2j}} \\ Y_{E2j} \end{pmatrix}, \begin{pmatrix} Y_{C2j} \\ Y_{E2j} \end{pmatrix} \sim_{i.i.d} N_2\left(\begin{pmatrix} \mu_3 = 6 \\ \mu_4 = 3 \end{pmatrix}, \begin{pmatrix} \sigma_3^2 = 2 & \sigma_{34} \\ \sigma_{34} & \sigma_4^2 = .5 \end{pmatrix}\right),$$

Bivariate lognormal:

$$\begin{pmatrix} C_{1j} \\ E_{1j} \end{pmatrix} \sim_{i.i.d} EXP\left[N_2\left(\begin{pmatrix} \mu_1 = 8 \\ \mu_2 = 4 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 = 2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 = .5 \end{pmatrix}\right)\right],$$
$$\begin{pmatrix} C_{2j} \\ E_{2j} \end{pmatrix} \sim_{i.i.d} EXP\left[N_2\left(\begin{pmatrix} \mu_3 = 6 \\ \mu_4 = 3 \end{pmatrix}, \begin{pmatrix} \sigma_3^2 = 2 & \sigma_{34} \\ \sigma_{34} & \sigma_4^2 = .5 \end{pmatrix}\right)\right].$$

[Tables 1 and 2 here]

The result of our simulation is presented in Tables 1 and 2. For data generated from normal distributions, all intervals give good coverages. Average interval lengths (presented in the parenthesis) are also comparable for all methods. For data generated from bivariate mixture and bivariate log-normal distributions, however, confidence intervals based on normal theory (Taylor's) are obviously inadequate. Fieller's intervals are also insufficient in term of coverage accuracy in some cases. Hinkley's intervals are very similar to Fieller's intervals. Overall, the bootstrap-t intervals (boott) appear best in term of coverage accuracy. However, as mentioned before, bootstrap-t intervals are often too wide. When facing with such wide intervals, one has to balance between the coverage accuracy and the precision of the estimate. Our new intervals give better coverage than Taylor's interval. They are comparable and sometimes better than the Fieller's and Hinkley's intervals. T_3 intervals have comparable coverage and are narrower than bootstrap-t intervals. We also considered other correlation structures and sample sizes. The results were similar and are not reported here.

It is clear from equation (3), $|q_1(x)|/\sqrt{N}$ plays an important factor in the expansion. In our simulation study for the 95% confidence interval, we saw that when $|q_1(1.96)|/\sqrt{N}$ was large (≥ 0.3), the coverages of normal based intervals were inadequate. Such intervals can be improved upon by the bootstrap-t or our new intervals T_3. On the other hand, when $|q_1(1.96)|/\sqrt{N}$ was small (< 0.3), normal based intervals gave reasonable coverage results.

In summary of our simulation in this section, we found the bootstrap-t intervals gave best coverage. Our intervals T_3 gave comparable coverage as bootstrap-t intervals, were about one-third narrower, could be computed easily (with a hand-held calculator), and also required less computing than the bootstrap-t intervals. T_3 intervals should be recommended for the ICER confidence intervals construction when data coming from skewed distributions and sample sizes are small.

4.2 *NHB Simulation*

We evaluated our three new intervals given in equation (8) against the normal interval, the bootstrap bias-corrected interval, and the bootstrap percentile-t interval. Details of these intervals are given in Appendix D.

We also simulated data from three setups: bivariate normal, bivariate lognormal, and bivariate mixture (costs are lognormal and effectiveness data are normal). With n_1, n_2, σ_{12}, and σ_{34} vary as presented in Tables 3 and 4, the data were generated as following:

Bivariate normal:

$$\begin{pmatrix} C_{1j} \\ E_{1j} \end{pmatrix} \sim_{i.i.d} N_2 \left(\begin{pmatrix} \mu_1 = 40,000 \\ \mu_2 = 60 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 = 200,000 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 = 5 \end{pmatrix} \right),$$

$$\begin{pmatrix} C_{2j} \\ E_{2j} \end{pmatrix} \sim_{i.i.d} N_2 \left(\begin{pmatrix} \mu_3 = 30,000 \\ \mu_4 = 50 \end{pmatrix}, \begin{pmatrix} \sigma_3^2 = 100,000 & \sigma_{34} \\ \sigma_{34} & \sigma_4^2 = 10 \end{pmatrix} \right),$$

Bivariate mixture:

$$\begin{pmatrix} C_{1j} \\ E_{1j} \end{pmatrix} = \begin{pmatrix} e^{Y_{C1j}} \\ Y_{E1j} \end{pmatrix}, \begin{pmatrix} Y_{C1j} \\ Y_{E1j} \end{pmatrix} \sim_{i.i.d} N_2 \left(\begin{pmatrix} \mu_1 = 8 \\ \mu_2 = 4 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 = 1 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 = 1 \end{pmatrix} \right),$$

$$\begin{pmatrix} C_{2j} \\ E_{2j} \end{pmatrix} = \begin{pmatrix} e^{Y_{C2j}} \\ Y_{E2j} \end{pmatrix}, \begin{pmatrix} Y_{C2j} \\ Y_{E2j} \end{pmatrix} \sim_{i.i.d} N_2 \left(\begin{pmatrix} \mu_3 = 6 \\ \mu_4 = 3 \end{pmatrix}, \begin{pmatrix} \sigma_3^2 = 2 & \sigma_{34} \\ \sigma_{34} & \sigma_4^2 = .5 \end{pmatrix} \right),$$

Bivariate lognormal:

$$\begin{pmatrix} C_{1j} \\ E_{1j} \end{pmatrix} \sim_{i.i.d} EXP \left[N_2 \left(\begin{pmatrix} \mu_1 = 8 \\ \mu_2 = 4 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 = 1 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 = 1 \end{pmatrix} \right) \right],$$

$$\begin{pmatrix} C_{2j} \\ E_{2j} \end{pmatrix} \sim_{i.i.d} EXP \left[N_2 \left(\begin{pmatrix} \mu_3 = 6 \\ \mu_4 = 3 \end{pmatrix}, \begin{pmatrix} \sigma_3^2 = 2 & \sigma_{34} \\ \sigma_{34} & \sigma_4^2 = .5 \end{pmatrix} \right) \right].$$

[Tables 3 and 4 here]

For data generated from normal and mixture distributions, all intervals give good coverage. Average interval length (in parenthesis) are also comparable. For data generated from lognormal distribution, however, intervals based on normal theory suffer deficiency in term of coverage accuracy. These coverages can be improved upon by the bootstrap-t and especially intervals based on T_3 transformation. T_3 intervals have several advantages over bootstrap methods including shorter interval lengths, easy to compute, and less computing time.

It is clear from equation (6) that coefficient $A = \frac{1}{Q^{3/2}}\left[\frac{B_1}{\nu^2} - \frac{B_2}{(1-\nu)^2}\right]$ in absolute value and sample sizes play an important role in the normal approximation. Our simulation suggests that when $|\hat{A}|/\sqrt{N}$ is small (≤ 0.3), normal intervals give good coverages. On the contrary, when $|\hat{A}|/\sqrt{N}$ is large (> 0.3), normal intervals suffer deficiency and can be improved by the bootstrap-t or the T_3 intervals. Thus, when dealing with skewed data, intervals based on T_3 transformation should be recommended.

5. Application

In this section, we applied the methods evaluated above to a real data set. In 2002, Katon et al. (2002) conducted a randomized study to assess whether a collaborative care intervention would increase the number of anxiety-free days for patients with panic disorder, compared to the usual primary care setting. The collaborative care intervention included a systematic patient education and approximately 2 visits with an on-site consulting psychiatrist. To demonstrate our methods, we consider total outpatient non-mental health costs and for measure of effectiveness, we consider the number of days a patient experiences panic disorder during the one-year study period.

The summary statistics are presented in Table 5. Distributions of costs and effectiveness are presented in Figure 2. Costs in both groups are highly skewed with coefficient of skewness 4.93 for the control group and 3.46 for the intervention group. On average, the control group incurred \$1181.95 and 77.1 days of panic attack more than the intervention group. The estimated ICER is 15.33 indicating that the intervention arm is dominant.

[Tables 5, 6, and Figure 2 here]

Confidence intervals for the ICER and the NHB are presented in Table 6. Both bootstrap-t and T_3 intervals are positive showing that the intervention is significantly dominant (the control group incurred more cost and more days of panic disorder). As anticipated, both Fieller and Hinkley intervals are similar and T_3 interval has shorter length than bootstrap-t. The estimated $\hat{q}(1.96)/\sqrt{N} = .9$, indicating the normal based interval is inadequate. Based on our simulation results, we would recommend using the T_3 interval as the confidence interval for the ICER.

A common willingness-to-pay is $\lambda = \$50,000$, using this value, the estimated NHB is 77.08. Confidence intervals for the NHB are presented in Table 6. All intervals are relatively similar, especially T_1, T_2, and normal intervals. The coefficient $\hat{A}/\sqrt{N}$ is 0.05 in this setting suggesting that the normal theory interval is adequate. All of these confidence intervals are strictly positive indicating, again, that intervention arm is cost-effective.

6. Discussion

In this paper, we derived the Edgeworth expansions for the studentized statistics for the ICER and the NHB and proposed new confidence intervals based on these Edgeworth expansion. We demonstrated via simulation studies that our methods have comparable coverage accuracy and are narrower compared with the current recommendation. In particular, for the ICER, when data were generated a from skewed distribution, our new intervals gave better coverages than the Taylor's interval. They are comparable and sometimes better than the Fieller's, and Hinkley's intervals. We found that Hinkley's method, that has not been adopted for the ICER, was similar to the Fieller's method in term of coverage accuracy and interval length. Intervals based on T_3 transformation were comparable to the bootstrap-t intervals in term of coverage but were about one-third narrower.

For the NHB, we saw that intervals based on T_3 transformation gave good coverages in all cases considered and they were comparable to the bootstrap-t intervals. However, our intervals were narrower than the bootstrap-t intervals and required less computing in term of bootstrap resampling.

When dealing with highly skewed data, our intervals based on the transformation T_3 should be recommended.

The remaining question is what one should choose between the ICER and the NHB. As it has been pointed out by other authors, and is summarized in our introduction section, each measure has its own advantages as well as disadvantages. The purpose of this paper is not to recommend one measure over the other, but to present new statistical methodology for both. We leave it to the readers to decide which measure is more appropriate for their works.

Acknowledgements

This work is supported in part by NIH grant AHRQ R01HS013105. The authors would like to thank Dr. N. David Yanez for pointing out the application of the Hinkley's method to the ICER and Dr. W. Katon for providing the data set used in this study. The opinions herein are solely those of the

authors and do not necessarily represent the authors′ institutions.

References

Briggs, A. and Fenn, P. (1998). Confidence intervals or surfaces? uncertainty on the cost-effectiveness plane. *Health Economics* **7**, 723–740.

Briggs, A., Mooney, C. and Wonderling, D. (1999). Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using monte carlo simulation. *Statistics in Medicine* **18**, 3245–3262.

Fan, M.-Y. and Zhou, X.-H. (2005). Constructing confidence intervals for incremental cost-effectiveness ratio. *Submitted* .

Fieller, E. (1954). Some problems in interval estimation. *Journal of the Royal Statistical Society, Ser B* **16**, 175–183.

Garber, A. and Phelps, C. (1997). Economic foundations of cost-effectiveness analysis. *Journal of Health Economics* **16**, 1–31.

Hall, P. (1992a). *The bootstrap and edgeworth expansion.* New York: Springer.

Hall, P. (1992b). On the removal of skewness by transformation. *Journal of the Royal Statistical Society. Ser. B* **54**, 221–228.

Hall, P. and Martin, M. (1988). On the bootstrap and two-sample problems. *Australian Journal of Statistics* **30A**, 179–192.

Hinkley, D. (1969). On the ratio of two correlated normal random variables. *Biometrika* **56**, 635–639.

Katon, W., Roy-Byrne, P., Russo, J. and Cowley, D. (2002). Cost-effectiveness and cost offset of a collaborative care intervention for primary care patients with panic disorder. *Archive of General Psychiatry* **59**, 1098–1104.

Stinnett, A. and Mullahy, J. (1998). Net health benefit: a new framework for the analysis of uncertainty in cost-effectiveness analysis. *Medical Decision Making* **18**, S68–S80.

Willan, A. (2001). Analysis, sample size, and power for estimating net health benefit from clinical trial data. *Control Clinical Trials* **22**, 228–237.

Willan, A. (2003). Analysing cost-effectiveness trials: net benefits. *in Statistical methods for cost-effectiveness research: a guide to current issues and future developments* (Briggs, A., ed.). London: OHE. pg. 8–23.

Zethraeus, N., Johannesson, M., Jonsson, B. and Lothgren, M. (2003). Advantages of using the net-benefit approach for analysing uncertainty in economic evaluation studies. *Pharmacoeconomics* **21**, 39–48.

Zhou, X.-H. and Dinh, P. (2005). Nonparametric confidence intervals for the one- and two-sample problems. *Biostatistics (in press)* .

Appendix A

Derivation of the ICER Edgeworth expansion

A1. Proof of theorem 1

Let $X = (X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, X_{10})$ and

$$Y_i = \frac{(C_{1i}-\mu_{C_1})}{\sigma_{C_1}},\ U_i = \frac{(E_{1i}-\mu_{E_1})}{\sigma_{E_1}},\ Z_i = \frac{(C_{2i}-\mu_{C_2})}{\sigma_{C_2}},\ W_i = \frac{(E_{2i}-\mu_{E_2})}{\sigma_{E_2}}$$

where

$$X_1 = \bar{Y},\quad X_2 = n_1^{-1}\sum_{i=1}^{n_1} Y_i^2,\quad X_3 = \bar{U},\quad X_4 = n_1^{-1}\sum_{i=1}^{n_1} U_i^2,$$
$$X_5 = \bar{Z},\quad X_6 = n_2^{-1}\sum_{i=1}^{n_2} Z_i^2,\quad X_7 = \bar{W},\quad X_8 = n_2^{-1}\sum_{i=1}^{n_2} W_i^2,$$
$$X_9 = n_1^{-1}\sum_{i=1}^{n_1} Y_iU_i,\quad X_{10} = n_2^{-1}\sum_{i=1}^{n_2} Z_iW_i.$$

Let $k(X) = \hat{\theta}_1 - \theta_1$, $h(X) = \hat{\sigma}_1^2$, $f(X) = \frac{k(X)}{h(X)^{1/2}}$, and $T_1 = N^{1/2} f(X)$. Let $V = E(X)$, then $f(V) = 0$. By Taylor expansion, we obtain,

$$f(X) = \frac{\partial f(V)}{\partial V}(X - V) + \frac{1}{2}(X - V)' \frac{\partial^2 f(V)}{\partial V^2}(X - V) + ...$$

Let $J_1 = N^{1/2}\left(\frac{\partial f(V)}{\partial V}(X - V) + \frac{1}{2}(X - V)' \frac{\partial^2 f(V)}{\partial V^2}(X - V)\right)$. Let A_1 and A_2 be given as in Theorem 1. The first three moments of J_1 are given by

$$EJ_1 = -(A_1 + A_2)N^{-1/2}, \;\; EJ_1^2 = 1 + O(N^{-1})$$

$$EJ_1^3 = -(7A_1 + 9A_2)N^{-1/2} + O(N^{-min(1,r+1/2)}).$$

Let K_{1N}, K_{2N}, K_{3N} be the first three cumulants of J_1. Then

$$K_{1N} = -(A_1 + A_2)N^{-1/2}, \;\; K_{2N} = 1 + O(N^{-1})$$

$$K_{3N} = -(4A_1 + 6A_2)N^{-1/2} + O(N^{-min(1,r+1/2)})$$

Let $\chi_N(t)$ be the characteristic function of J_1. Applying Taylor's expansion to function $f(x) = e^x$, we obtain

$$\chi_N(t) = exp(-\frac{t^2}{2})\left[1 + N^{-1/2}\left(-(A_1 + A_2)(it) - \frac{(4A_1+6A_2)}{6}(it)^3\right) + O(N^{-min(1,r+1/2)}\right]$$

Let $r_1(it) = -(A_1 + A_2)(it) - \frac{(4A_1+6A_2)}{6}(it)^3$, we can write

$$\chi_N(t) = exp(-\frac{t^2}{2})\left[1 + N^{-1/2} r_1(it) + O(N^{-min(1,r+1/2)})\right]$$

Since $\chi_N(t) = \int e^{itx} dP(J_1 \leq x)$ and $e^{-t^2/2} = \int e^{itx} d\Phi(x)$, this suggests that

$$P(J_1 \leq x) = \Phi(x) + N^{-1/2} R_1(x) + O(N^{-min(1,r+1/2)})$$

where $R_1(x)$ is such a function that its Fourier-Stieltjes transform equals to $r_1(it)e^{-t^2/2}$. Applying integration by part, we obtain,

$$R_1(x) = [(A_1 + A_2) + \frac{4A_1+6A_2}{6}(x^2 - 1)]\phi(x).$$

Therefore,

$$P(J_1 \leq x) = \Phi(x) + N^{-1/2} q_1(x)\phi(x) + O(N^{-min(1,r+1/2)}).$$

Since $T_1 = J_1 + O(N^{-1})$, Theorem 1 follows.

A2. Confidence intervals construction

Let $c_1 = -(A_1 + A_2)$ and $c_2 = -(7A_1 + 9A_2)$. Following Hall (1992a), let $T_1' = T_1 - N^{-1/2}\hat{c}_1$, then

$$P(T_1' \leq x) = P(T_1 - N^{-1/2}\hat{c}_1 \leq x) = P(T_1 \leq x + N^{-1/2}\hat{c}_1)$$
$$= \Phi(x+N^{-1/2}\hat{c}_1)+N^{-1/2}q_1(x+N^{-1/2}\hat{c}_1)\phi(x+N^{-1/2}\hat{c}_1)+O(N^{-min(1,r+1/2)})$$

Applying Taylor expansion, with $\gamma = \frac{2A_1+3A_2}{3}$, and $\hat{\gamma}$ is its estimate, we get

$$P(T_1' \leq x) = \Phi(x) + N^{-1/2}\hat{\gamma}(x^2-1)\phi(x) + O(N^{-min(1,r+1/2)})$$

Similar to Zhou and Dinh (2005), intervals (4) follow.

Appendix B

Proof of Theorem 2

Let X be defined as in the proof of Theorem 1, $k(X) = \hat{\theta}_2 - \theta_2$, $h(X) = \hat{\sigma}_2^2$, $f(X) = \frac{k(X)}{h(X)^{1/2}}$, and $T_2 = N^{1/2}f(X)$. Let $V = E(X)$, then $f(V) = 0$. By Taylor expansion, we obtain,

$$f(X) = \frac{\partial f(V)}{\partial V}(X-V) + \frac{1}{2}(X-V)'\frac{\partial^2 f(V)}{\partial V^2}(X-V) + ...$$

Let $J_2 = N^{1/2}\left(\frac{\partial f(V)}{\partial V}(X-V) + \frac{1}{2}(X-V)'\frac{\partial^2 f(V)}{\partial V^2}(X-V)\right)$. With coefficient A given in Theorem 2, the first 3 moments of J_2 are:

$$EJ_2 = -\tfrac{1}{2}AN^{-1/2}, \quad EJ_2^2 = 1+O(N^{-1}), \quad EJ_2^3 = -\tfrac{7}{2}AN^{-1/2}+O(N^{-min(1,r+1/2)}).$$

Let K_{1N}, K_{2N}, K_{3N} be the first three cumulants of J_2. Then,

$$K_{1N} = -\tfrac{1}{2}AN^{-1/2}+O(N^{-min(1,r+1/2)}), \quad K_{2N} = 1+O(N^{-min(1,r+1/2)}),$$
$$K_{3N} = -2AN^{-1/2} + O(N^{-min(1,r+1/2)}).$$

Let $\chi_N(t)$ be the characteristic function of J_2. Applying Taylor expansion to an exponential function, we obtain

$$\chi_N(t) = exp(-\tfrac{t^2}{2})\left\{1 + N^{-1/2}(-\tfrac{1}{2}A(it) - \tfrac{2A}{6}(it)^3) + O(N^{-min(1,r+1/2)})\right\}.$$

Letting $R_2(it) = (-\frac{1}{2}A(it) - \frac{2A}{6}(it)^3)$, we can write

$$\chi_N(t) = exp(-\frac{t^2}{2})\left\{1 + N^{-1/2}R_2(it) + O(N^{-min(1,r+1/2)})\right\}.(*)$$

Since $\chi_N(t) = \int_{-\infty}^{\infty} e^{itx}dp(J_2 \leq x)$ and $e^{-t^2/2} = \int_{-\infty}^{\infty} e^{itx}d\Phi(x)$, expression (*) suggests that

$$P(J_2 \leq x) = \Phi(x) + N^{-1/2}R_2(X) + O(N^{-min(1,r+1/2)})$$

where $R_2(X)$ is such a function that its Fourier-Stieltjes transform equals to $r_2(it)e^{-t^2/2}$. Applying integration by part, we obtain

$$R_2(x) = \frac{A}{2} + \frac{2A}{6}(x^2 - 1)\phi(x) = \frac{A}{6}(2x^2 + 1)\phi(x).$$

Therefore,

$$P(J_2 \leq x) = \Phi(x) + N^{-1/2}q_2(x)\phi(x) + O(N^{-min(1,r+1/2)}).$$

Since $T_2 = J_2 + O(N^{-1})$, Theorem 2 follows.

Appendix C

Details of Methods for the ICER Simulation

C1. Taylor's interval

The variance of the $\widehat{ICER}$ can be estimated as in equation (2). The asymptotic confidence interval (namely Taylor's interval) can be derived using the central limit theorem and is given by $(\widehat{ICER} - z_{1-\alpha/2}\hat{\sigma}_1,\ \widehat{ICER} + z_{1-\alpha/2}\hat{\sigma}_1)$.

C2. Fieller's interval

The Fieller (Fieller, 1954) method assumes that cost and effectiveness follow a bivariate normal distribution. Thus, $\hat{\Delta}_2 - R\hat{\Delta}_1$ is normally distributed with zero mean, with R denotes the ICER and $\hat{\Delta}_1 = \bar{E}_1 - \bar{E}_2$, and $\hat{\Delta}_2 = \bar{C}_1 - \bar{C}_2$. Consequently,

$$\frac{\hat{\Delta}_2 - R\hat{\Delta}_1}{\sqrt{\widehat{Var}(\hat{\Delta}_2 - R\hat{\Delta}_1)}} \sim N(0, 1).$$

The $(1-\alpha)100\%$ confidence interval for the ICER (denoted Fieller) can be obtained by equating the formula

$$\frac{\hat{\Delta}_2 - R\hat{\Delta}_1}{\sqrt{\widehat{Var}(\hat{\Delta}_2 - R\hat{\Delta}_1)}} = z_{1-\alpha/2}$$

and solve for R. This is equivalent to solving the quadratic equation

$$XR^2 - 2YR + Z = 0 \tag{C.1}$$

where

$$X = (\bar{E}_1 - \bar{E}_2)^2 - z^2_{1-\alpha/2}\left[\frac{S^2_{E_1}}{n_1} + \frac{S^2_{E_2}}{n_2}\right],$$
$$Y = (\bar{E}_1 - \bar{E}_2)(\bar{C}_1 - \bar{C}_2) - z^2_{1-\alpha/2}\left[\frac{S_{E_1C_1}}{n_1} + \frac{S_{E_2C_2}}{n_2}\right],$$
$$Z = (\bar{C}_1 - \bar{C}_2)^2 - z^2_{1-\alpha/2}\left[\frac{S^2_{C_1}}{n_1} + \frac{S^2_{C_2}}{n_2}\right].$$

Confidence limits for the ICER are the solutions of the equation (C.1).

C3. Bootstrap-t interval

The t-statistic $T = (\widehat{ICER} - ICER)/\hat{\sigma}_1$ is normally distributed based on the central limit theorem when the sample size is large. When this assumption is violated, an alternative is to approximate its distribution by a bootstrap method, namely the bootstrap-t. For each bootstrap sample b, the bootstrap-t computes another t-statistic $T^*_b = (\widehat{ICER}^*_b - \widehat{ICER})/\hat{\sigma}^*_{1_b}$. These t-statistics will then be sorted to find the $\alpha/2$ and $(1-\alpha/2)$ percentiles, denoted by $\hat{t}^{(\alpha/2)}$ and $\hat{t}^{(1-\alpha/2)}$. The resulting $(1-\alpha)100\%$ bootstrap-t interval (denoted boott) is $(\widehat{ICER} - \hat{t}^{(1-\alpha/2)}\hat{\sigma}_1,\ \ \widehat{ICER} - \hat{t}^{(\alpha/2)}\hat{\sigma}_1)$.

C4. Hinkley's interval

The Hinkley (Hinkley, 1969) method also assumes the numerator and denominator of the ICER come from a bivariate normal distribution. The distribution function $F(w)$ of the ICER is given by (Hinkley, 1969),

$$\begin{aligned} F(w) &= L\left(\frac{\Delta_2-\Delta_1 w}{\sqrt{\Gamma_1\Gamma_2}a(w)/N}, -\frac{\Delta_1}{\sqrt{\Gamma_1/N}}; \frac{w\sqrt{\Gamma_1/N}-\rho\sqrt{\Gamma_2/N}}{\sqrt{\Gamma_1\Gamma_2}a(w)/N}\right) \\ &\quad +L\left(\frac{\Delta_1 w-\Delta_2}{\sqrt{\Gamma_1\Gamma_2}a(w)/N}, \frac{\Delta_1}{\sqrt{\Gamma_1/N}}; \frac{w\sqrt{\Gamma_1/N}-\rho\sqrt{\Gamma_2/N}}{\sqrt{\Gamma_1\Gamma_2}a(w)/N}\right) \end{aligned}$$

where, $\Delta_1, \Delta_2, \Gamma_1, \Gamma_2, \Gamma_{12}$ are given in Theorem 1, and

$$L(h,k;\gamma) = \frac{1}{2\pi\sqrt{1-\gamma^2}}\int_h^\infty\int_k^\infty exp\left[-\frac{x^2-2\gamma xy+y^2}{2(1-\gamma^2)}\right]dxdy,$$

$$a(w) = \left(\frac{w^2}{\Gamma_2/N} - \frac{2\rho w}{\sqrt{\Gamma_1\Gamma_2}/N} + \frac{1}{\Gamma_1/N}\right)^{1/2}, \;\; \rho = \Gamma_{12}/\sqrt{\Gamma_1\Gamma_2}.$$

The $(1-\alpha)100\%$ confidence interval (denoted Hinkley) for the ICER is given by (w_1, w_2) where $F(w_1) = \alpha/2, \;\; F(w_2) = 1-\alpha/2$.

C5. Three new transformational intervals

The three new transformational intervals based on Edgeworth expansion are given in equation (4). They are denoted by T_1, T_2, and T_3 in our simulation.

Appendix D

Details of Methods for the NHB Simulation

D1. Normal interval

The asymptotic variance of the NHB can be estimated as in equation (5). By the central limit theorem, the asymptotic normal confidence interval (denoted normal) for the NHB is given by $\left(\widehat{NHB} - z_{1-\alpha/2}\hat{\sigma}_2, \quad \widehat{NHB} + z_{1-\alpha/2}\hat{\sigma}_2\right)$.

D2. BCa interval

For each bootstrap sample b, the estimate $\widehat{NHB}_b^*$ is computed. The bootstrap replicates are then ordered from smallest to largest. The bias-corrected acceleration (BCa) confidence interval is given by $(\widehat{NHB}^{*(\alpha_1)},\ \widehat{NHB}^{*(\alpha_2)})$ where

$$\alpha_1 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z_\alpha}{1 - \hat{\alpha}(\hat{z}_0 + z_\alpha)}\right),\ \alpha_2 = \Phi\left(\hat{z}_0 + \frac{\hat{z}_0 + z_{1-\alpha}}{1 - \hat{a}(\hat{z}_0 + z_{1-\alpha})}\right),$$
$$\hat{z_0} = \Phi^{-1}\left(\frac{\#\{\widehat{NHB}^*(b) < \widehat{NHB}\}}{B}\right).$$

The acceleration constant $\hat{a} = \frac{1}{6}\hat{\sigma}_2^{-3}(n_1^{-2}\hat{\gamma}_1 - n_2^{-2}\hat{\gamma}_2)$ is given in Hall and Martin (1988) where

$$\hat{\gamma}_1 = \frac{1}{n_1}\sum_{j=1}^{n_1}\left[(E_{1j} - \bar{E}_1) - \frac{(C_{1j} - \bar{C}_1)}{\lambda}\right]^3,\ \hat{\gamma}_2 = \frac{1}{n_2}\sum_{j=1}^{n_2}\left[(E_{2j} - \bar{E}_2) - \frac{(C_{2j} - \bar{C}_2)}{\lambda}\right]^3.$$

D3. Bootstrap-t interval

As described previously, For each bootstrap sample b, the bootstrap-t computes another t-statistic $T_b^* = (\widehat{NHB}_b^* - \widehat{NHB})/\hat{\sigma}_{2_b}^*$. These t-statistics will then be sorted to find the $\alpha/2$ and $(1-\alpha/2)$ percentiles, denoted by $\hat{t}^{(\alpha/2)}$ and $\hat{t}^{(1-\alpha/2)}$. The resulting $(1-\alpha)100\%$ bootstrap-t interval (denoted boott) is $(\widehat{NHB} - \hat{t}^{(1-\alpha/2)}\hat{\sigma}_2,\ \widehat{NHB} - \hat{t}^{(\alpha/2)}\hat{\sigma}_2)$.

D4. Three new transformational intervals

The three new transformational intervals based on Edgeworth expansion are given in equation (8). They are denoted by T_1, T_2, and T_3 in our simulation.

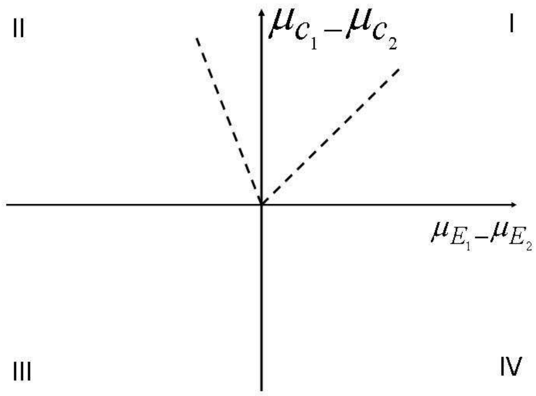

Figure 1. The $\Delta E - \Delta C$ Plane

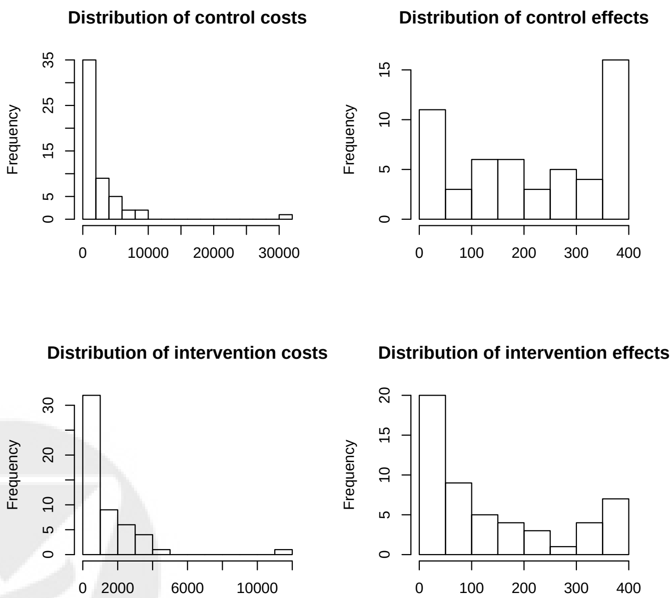

Figure 2. Distributions of costs and effectiveness of two groups

Table 1

Coverage of 95% 2-sided confidence intervals for the ICER for bivariate normal and mixture data ($n_1 = n_2 = 50$)

σ_{12}	σ_{34}	Taylor	Fieller	Hinkley	Boott	T_1	T_2	T_3	$\frac{\lvert q_1(1.96)\rvert}{\sqrt{N}}$
Normal									
800	800	0.9513 (194.85)	0.9505 (196.25)	0.9504 (196.24)	0.9502 (197.83)	0.9499 (196.79)	0.9502 (194.65)	0.9526 (205.23)	0.210
800	200	0.9476 (204.50)	0.9462 (205.72)	0.9460 (205.72)	0.9463 (207.49)	0.9454 (206.24)	0.9474 (204.24)	0.9497 (215.20)	0.209
200	800	0.9452 (204.43)	0.9463 (205.77)	0.9463 (205.76)	0.9469 (207.13)	0.9465 (206.18)	0.9453 (204.22)	0.9515 (215.15)	0.208
200	200	0.9480 (213.25)	0.9489 (214.64)	0.9489 (214.64)	0.9500 (216.15)	0.9499 (215.19)	0.9481 (213.01)	0.9519 (224.52)	0.208
-800	-800	0.9469 (240.41)	0.9476 (241.98)	0.9475 (241.99)	0.9481 (243.63)	0.9474 (242.64)	0.9469 (240.17)	0.9532 (253.11)	0.209
-800	-200	0.9479 (232.56)	0.9485 (233.99)	0.9484 (234.00)	0.9495 (235.68)	0.9485 (234.56)	0.9478 (232.37)	0.9541 (244.82)	0.209
-200	-800	0.9489 (233.34)	0.9473 (234.76)	0.9473 (234.77)	0.9477 (236.31)	0.9475 (235.37)	0.9487 (233.11)	0.9486 (245.53)	0.209
-200	-200	0.9497 (225.12)	0.9477 (226.57)	0.9478 (226.57)	0.9486 (228.07)	0.9477 (227.12)	0.9494 (224.92)	0.9535 (236.96)	0.209
Mixture									
0.9	0.9	0.8215 (8648.77)	0.8296 (9013.98)	0.8297 (9013.17)	0.8877 (16360.3)	0.8586 (12206.9)	0.8286 (8548.10)	0.8675 (8586.88)	0.588
0.2	0.2	0.8484 (10358.3)	0.8723 (10730.5)	0.8722 (10730.1)	0.8987 (16547.7)	0.8885 (15160.4)	0.8564 (10230.8)	0.8871 (10294.3)	0.680
0.9	0.2	0.8339 (8826.95)	0.8495 (9170.84)	0.8496 (9170.13)	0.8835 (15647.7)	0.8606 (12226.8)	0.8405 (8730.71)	0.8815 (8753.06)	0.566
0.2	0.9	0.8440 (10379.6)	0.8708 (10719.0)	0.8709 (10718.9)	0.9050 (17435.9)	0.8905 (15062.8)	0.8527 (10241.8)	0.8905 (10249.0)	0.693
-0.9	-0.9	0.8580 (13329.0)	0.8980 (13934.3)	0.8980 (13928.6)	0.8979 (19227.1)	0.9136 (21919.5)	0.8665 (13132.9)	0.9013 (13142.8)	0.850
-0.2	-0.2	0.8638 (11683.8)	0.8966 (12118.9)	0.8966 (12118.1)	0.9105 (19181.7)	0.9069 (17949.6)	0.8727 (11529.3)	0.9051 (11557.3)	0.747
-0.9	-0.2	0.8563 (13236.9)	0.8955 (13837.8)	0.8955 (13832.7)	0.8957 (19861.9)	0.9072 (21737.3)	0.8668 (13022.3)	0.9009 (13021.4)	0.857
-0.2	-0.9	0.8665 (11794.2)	0.8995 (12238.4)	0.8996 (12236.5)	0.9068 (18771.1)	0.9060 (17881.2)	0.8750 (11642.6)	0.9040 (11709.9)	0.737

Table 2

Coverage of 95% 2-sided confidence intervals for the ICER for bivariate lognormal data

n_1	n_2	σ_{12}	σ_{34}	Taylor	Fieller	Hinkley	Boott	T_1	T_2	T_3	$\frac{\lvert q_1(1.96)\rvert}{\sqrt{N}}$
50	50	0.9	0.9	0.8086 (140.85)	0.8129 (159.46)	0.8131 (157.71)	0.8678 (264.39)	0.8353 (167.08)	0.8142 (139.37)	0.8628 (140.73)	0.410
50	50	0.2	0.2	0.8556 (247.53)	0.9037 (266.87)	0.9037 (265.69)	0.9030 (385.72)	0.8961 (363.12)	0.8609 (244.59)	0.8944 (245.40)	0.684
50	50	0.9	0.2	0.8361 (145.22)	0.8482 (280.12)	0.8486 (161.80)	0.8740 (241.04)	0.8478 (169.51)	0.8389 (143.99)	0.8855 (145.36)	0.393
50	50	0.2	0.9	0.8620 (244.72)	0.9065 (262.02)	0.9065 (261.80)	0.9095 (396.32)	0.8974 (362.76)	0.8683 (241.74)	0.8991 (241.85)	0.694
50	50	-0.9	-0.9	0.8698 (314.32)	0.9242 (344.01)	0.9242 (343.53)	0.9120 (452.35)	0.9337 (540.66)	0.8785 (309.79)	0.9111 (310.57)	0.854
50	50	-0.2	-0.2	0.8685 (283.35)	0.9179 (307.39)	0.9182 (307.07)	0.9148 (431.87)	0.9169 (445.67)	0.8760 (279.69)	0.9076 (280.49)	0.772
50	50	-0.9	-0.2	0.8690 (315.91)	0.9191 (348.47)	0.9190 (345.30)	0.9063 (471.17)	0.9287 (534.66)	0.8770 (311.41)	0.9068 (312.27)	0.856
50	50	-0.2	-0.9	0.8649 (283.50)	0.9142 (306.44)	0.9142 (306.18)	0.9138 (428.89)	0.9180 (447.68)	0.8714 (279.99)	0.9028 (279.90)	0.773
100	100	0.9	0.9	0.8427 (108.31)	0.8416 (114.17)	0.8416 (114.17)	0.8944 (176.64)	0.8643 (135.07)	0.8482 (107.31)	0.8776 (106.65)	0.463
100	100	0.2	0.2	0.8829 (177.37)	0.9140 (182.27)	0.9139 (182.27)	0.9210 (242.29)	0.9067 (236.14)	0.8861 (176.16)	0.9088 (175.16)	0.594
100	100	0.9	0.2	0.8583 (111.36)	0.8636 (116.91)	0.8636 (116.90)	0.8956 (170.36)	0.8731 (136.81)	0.8637 (110.53)	0.8937 (110.10)	0.424
100	100	0.2	0.9	0.8823 (177.06)	0.9098 (181.67)	0.9097 (181.68)	0.9186 (257.28)	0.9030 (238.88)	0.8858 (175.56)	0.9046 (174.44)	0.603
100	100	-0.9	-0.9	0.8899 (216.13)	0.9254 (222.38)	0.9255 (222.37)	0.9268 (268.00)	0.9264 (302.63)	0.8946 (214.27)	0.9144 (212.91)	0.684
100	100	-0.2	-0.2	0.8912 (197.24)	0.9231 (203.04)	0.9232 (203.07)	0.9268 (270.97)	0.9198 (269.52)	0.8959 (195.70)	0.9146 (194.49)	0.640
100	100	-0.9	-0.2	0.8977 (215.06)	0.9268 (221.96)	0.9269 (221.96)	0.9258 (270.79)	0.9272 (301.68)	0.9017 (213.40)	0.9186 (212.14)	0.689
100	100	-0.2	-0.9	0.8975 (200.30)	0.9250 (206.54)	0.9250 (206.53)	0.9269 (268.74)	0.9186 (272.52)	0.9017 (198.76)	0.9183 (198.05)	0.643
50	100	0.9	0.9	0.8021 (136.73)	0.8024 (152.29)	0.8023 (152.20)	0.8757 (264.07)	0.8325 (161.19)	0.8081 (135.35)	0.8452 (134.71)	0.417
50	100	0.2	0.2	0.8520 (242.38)	0.8980 (256.34)	0.8979 (256.30)	0.9059 (394.79)	0.8957 (355.36)	0.8557 (239.85)	0.8829 (238.10)	0.680
50	100	0.9	0.2	0.8166 (141.31)	0.8216 (158.12)	0.8216 (157.05)	0.8747 (257.89)	0.8411 (163.91)	0.8229 (140.26)	0.8558 (139.71)	0.394
50	100	0.2	0.9	0.8565 (243.40)	0.9013 (258.20)	0.9013 (258.11)	0.9130 (406.37)	0.8967 (355.44)	0.8614 (240.98)	0.8846 (239.44)	0.684
50	100	-0.9	0.9	0.8672 (307.95)	0.9190 (331.06)	0.9190 (330.95)	0.9165 (456.05)	0.9296 (521.98)	0.8746 (304.18)	0.8984 (302.28)	0.849
50	100	-0.2	0.2	0.8628 (275.45)	0.9105 (294.44)	0.9105 (294.36)	0.9174 (443.17)	0.9176 (429.66)	0.8691 (272.26)	0.8907 (270.54)	0.768
50	100	-0.9	0.2	0.8638 (312.66)	0.9148 (335.71)	0.9146 (335.63)	0.9146 (492.50)	0.9282 (525.66)	0.8711 (308.67)	0.8954 (305.64)	0.853
50	100	-0.2	0.9	0.8649 (272.82)	0.9144 (290.59)	0.9144 (290.28)	0.9211 (426.69)	0.9168 (421.22)	0.8717 (269.80)	0.8942 (267.85)	0.768
100	50	0.9	0.9	0.8520 (111.28)	0.8524 (118.17)	0.8525 (118.23)	0.8929 (182.74)	0.8664 (139.71)	0.8578 (110.38)	0.8883 (110.47)	0.454
100	50	0.2	0.2	0.8875 (183.21)	0.9171 (190.09)	0.9172 (190.27)	0.9171 (243.31)	0.9084 (244.94)	0.8924 (181.59)	0.9175 (181.67)	0.596
100	50	0.9	0.2	0.8753 (118.64)	0.8813 (125.07)	0.8812 (125.06)	0.8961 (173.17)	0.8786 (145.78)	0.8772 (117.83)	0.9071 (118.34)	0.417
100	50	0.2	0.9	0.8838 (176.32)	0.9118 (182.50)	0.9119 (182.56)	0.9158 (239.80)	0.9059 (237.60)	0.8887 (174.73)	0.9099 (174.60)	0.599
100	50	-0.9	-0.9	0.9000 (225.19)	0.9296 (235.43)	0.9296 (235.42)	0.9175 (295.21)	0.9219 (322.09)	0.9045 (223.31)	0.9244 (223.25)	0.691
100	50	-0.2	-0.2	0.8918 (204.59)	0.9271 (212.21)	0.9271 (212.07)	0.9182 (262.34)	0.9167 (284.15)	0.8962 (202.72)	0.9220 (202.21)	0.647
100	50	-0.9	-0.2	0.8968 (225.22)	0.9295 (234.12)	0.9295 (234.12)	0.9196 (285.79)	0.9232 (321.77)	0.9020 (223.32)	0.9234 (222.23)	0.693
100	50	-0.2	-0.9	0.8988 (205.02)	0.9286 (213.28)	0.9285 (213.20)	0.9201 (262.86)	0.9158 (281.68)	0.9027 (203.38)	0.9239 (202.77)	0.642

Table 3

Coverage of 95% 2-sided confidence intervals for the NHB for bivariate normal and mixture data ($n_1 = n_2 = n$)

n	σ_{12}	σ_{34}	Normal	Boott	BCa	T_1	T_2	T_3	$\lvert\hat{A}\rvert/\sqrt{N}$
Normal									
50	800	800	0.9468 (2.14)	0.9500 (2.18)	0.9426 (2.11)	0.9467 (2.14)	0.9469 (2.14)	0.9439 (2.30)	0.0001
50	800	200	0.9486 (2.14)	0.9512 (2.18)	0.9445 (2.12)	0.9479 (2.14)	0.9485 (2.14)	0.9455 (2.30)	0.0002
50	200	800	0.9465 (2.14)	0.9498 (2.18)	0.9428 (2.12)	0.9469 (2.14)	0.9464 (2.14)	0.9416 (2.30)	0.0002
50	200	200	0.9457 (2.14)	0.9493 (2.18)	0.9447 (2.12)	0.9459 (2.14)	0.9458 (2.14)	0.9440 (2.30)	0.0001
50	-800	-800	0.9495 (2.15)	0.9512 (2.19)	0.9463 (2.12)	0.9496 (2.15)	0.9496 (2.15)	0.9468 (2.31)	0.0003
50	-800	-200	0.9490 (2.14)	0.9530 (2.18)	0.9465 (2.12)	0.9490 (2.15)	0.9490 (2.14)	0.9453 (2.31)	0.0000
50	-200	-800	0.9480 (2.14)	0.9509 (2.18)	0.9436 (2.12)	0.9483 (2.14)	0.9474 (2.14)	0.9441 (2.31)	0.0004
50	-200	-200	0.9471 (2.14)	0.9496 (2.18)	0.9448 (2.12)	0.9475 (2.14)	0.9471 (2.14)	0.9447 (2.31)	0.0000
Mixture									
50	-0.8	0.8	0.9494 (0.71)	0.9514 (0.72)	0.9446 (0.70)	0.9507 (0.71)	0.9495 (0.71)	0.9445 (0.76)	0.014
50	-0.2	0.2	0.9438 (0.69)	0.9480 (0.70)	0.9411 (0.68)	0.9441 (0.69)	0.9437 (0.69)	0.9452 (0.74)	0.002
50	-0.8	0.2	0.9516 (0.71)	0.9539 (0.73)	0.9459 (0.71)	0.9520 (0.72)	0.9517 (0.71)	0.9457 (0.77)	0.015
50	-0.2	0.8	0.9461 (0.68)	0.9503 (0.70)	0.9430 (0.68)	0.9460 (0.68)	0.9462 (0.68)	0.9435 (0.74)	0.001
50	0.8	-0.8	0.9476 (0.65)	0.9508 (0.67)	0.9447 (0.65)	0.9483 (0.65)	0.9477 (0.65)	0.9414 (0.70)	0.005
50	0.2	-0.2	0.9472 (0.68)	0.9506 (0.69)	0.9432 (0.67)	0.9471 (0.68)	0.9469 (0.68)	0.9402 (0.73)	0.001
50	0.8	-0.2	0.9474 (0.65)	0.9515 (0.66)	0.9427 (0.64)	0.9468 (0.65)	0.9475 (0.65)	0.9444 (0.70)	0.009
50	0.2	-0.8	0.9492 (0.68)	0.9518 (0.69)	0.9442 (0.67)	0.9499 (0.68)	0.9496 (0.68)	0.9451 (0.73)	0.004

Table 4

Coverage of 95% 2-sided confidence intervals for the NHB for bivariate lognormal data

n_1	n_2	σ_{12}	σ_{34}	Normal	Boott	BCa	T_1	T_2	T_3	$\lvert\hat{A}\rvert/\sqrt{N}$
50	50	0.8	0.8	0.9025 (60.88)	0.9262 (76.77)	0.9052 (64.40)	0.8768 (85.87)	0.8971 (61.06)	0.9355 (64.61)	0.332
50	50	0.2	0.2	0.8997 (60.95)	0.9277 (77.09)	0.9059 (64.41)	0.8727 (85.80)	0.8942 (61.13)	0.9357 (64.69)	0.331
50	50	0.8	0.2	0.8929 (60.88)	0.9197 (76.94)	0.8991 (64.39)	0.8661 (86.31)	0.8881 (61.07)	0.9264 (64.61)	0.333
50	50	0.2	0.8	0.8983 (60.81)	0.9272 (77.22)	0.9041 (64.35)	0.8664 (85.99)	0.8923 (60.99)	0.9327 (64.53)	0.332
50	50	-0.8	-0.8	0.9005 (61.05)	0.9294 (77.31)	0.9052 (64.62)	0.8716 (86.03)	0.8943 (61.23)	0.9336 (64.78)	0.333
50	50	-0.2	-0.2	0.9082 (61.25)	0.9308 (77.52)	0.9106 (64.79)	0.8792 (86.34)	0.9024 (61.43)	0.9401 (65.00)	0.332
50	50	-0.8	-0.2	0.9047 (61.07)	0.9279 (77.31)	0.9059 (64.67)	0.8728 (86.41)	0.8969 (61.25)	0.9351 (64.80)	0.331
50	50	-0.2	-0.8	0.8965 (61.05)	0.9229 (77.24)	0.9027 (64.57)	0.8704 (86.23)	0.8913 (61.23)	0.9304 (64.79)	0.331
100	100	-0.8	0.8	0.9192 (44.42)	0.9338 (51.20)	0.9180 (46.43)	0.8968 (56.67)	0.9145 (44.48)	0.9354 (45.59)	0.294
100	100	-0.2	0.2	0.9207 (44.65)	0.9352 (51.75)	0.9186 (46.74)	0.8981 (57.28)	0.9178 (44.72)	0.9367 (45.83)	0.296
100	100	-0.8	0.2	0.9230 (44.87)	0.9330 (52.63)	0.9200 (47.02)	0.9011 (57.76)	0.9196 (44.94)	0.9374 (46.05)	0.295
100	100	-0.2	0.8	0.9188 (44.66)	0.9354 (51.90)	0.9187 (46.72)	0.8951 (57.22)	0.9151 (44.73)	0.9354 (45.83)	0.297
100	100	0.8	-0.8	0.9169 (44.37)	0.9367 (51.53)	0.9208 (46.41)	0.8921 (56.85)	0.9117 (44.44)	0.9377 (45.54)	0.295
100	100	0.2	-0.2	0.9185 (44.73)	0.9344 (52.13)	0.9183 (46.81)	0.8971 (57.48)	0.9150 (44.80)	0.9378 (45.90)	0.296
100	100	0.8	-0.2	0.9185 (44.32)	0.9363 (51.30)	0.9215 (46.36)	0.8955 (56.76)	0.9145 (44.38)	0.9378 (45.48)	0.294
100	100	0.2	-0.8	0.9191 (44.59)	0.9319 (51.68)	0.9167 (46.65)	0.8939 (57.34)	0.9153 (44.66)	0.9377 (45.76)	0.296
50	100	0.8	0.8	0.9007 (60.80)	0.9292 (78.83)	0.9066 (64.56)	0.8671 (87.33)	0.8945 (60.95)	0.9234 (62.97)	0.345
50	100	0.2	0.2	0.8947 (60.40)	0.9290 (77.50)	0.9033 (64.11)	0.8625 (86.63)	0.8887 (60.54)	0.9241 (62.55)	0.344
50	100	0.8	0.2	0.8945 (60.35)	0.9260 (77.26)	0.8999 (64.02)	0.8631 (86.71)	0.8894 (60.49)	0.9180 (62.50)	0.346
50	100	0.2	0.8	0.8930 (60.21)	0.9272 (77.03)	0.9021 (63.85)	0.8612 (86.55)	0.8864 (60.35)	0.9204 (62.35)	0.347
50	100	-0.8	-0.8	0.8966 (60.87)	0.9280 (78.80)	0.9040 (64.72)	0.8691 (87.30)	0.8913 (61.01)	0.9198 (63.03)	0.346
50	100	-0.2	-0.2	0.9035 (61.33)	0.9355 (79.30)	0.9115 (65.20)	0.8701 (88.56)	0.8965 (61.47)	0.9280 (63.50)	0.349
50	100	-0.8	-0.2	0.9005 (60.36)	0.9310 (77.31)	0.9082 (64.17)	0.8712 (86.40)	0.8968 (60.50)	0.9246 (62.51)	0.346
50	100	-0.2	-0.8	0.8892 (60.77)	0.9266 (79.15)	0.9023 (64.59)	0.8569 (88.00)	0.8816 (60.92)	0.9186 (62.93)	0.347
100	50	-0.8	0.8	0.9202 (45.19)	0.9320 (51.53)	0.9170 (46.98)	0.9012 (56.46)	0.9179 (45.26)	0.9432 (46.92)	0.274
100	50	-0.2	0.2	0.9212 (45.42)	0.9322 (51.97)	0.9177 (47.34)	0.9006 (57.34)	0.9181 (45.50)	0.9436 (47.16)	0.276
100	50	-0.8	0.2	0.9192 (45.55)	0.9269 (52.20)	0.9142 (47.51)	0.9008 (57.61)	0.9159 (45.63)	0.9388 (47.29)	0.278
100	50	-0.2	0.8	0.9233 (45.18)	0.9289 (51.43)	0.9162 (46.97)	0.9084 (56.38)	0.9204 (45.25)	0.9408 (46.91)	0.273
100	50	0.8	-0.8	0.9221 (45.16)	0.9304 (51.49)	0.9177 (47.04)	0.9019 (56.61)	0.9198 (45.23)	0.9432 (46.89)	0.275
100	50	0.2	-0.2	0.9239 (45.50)	0.9323 (51.49)	0.9186 (47.47)	0.9053 (57.60)	0.9213 (45.57)	0.9442 (47.23)	0.277
100	50	0.8	-0.2	0.9230 (45.35)	0.9349 (51.49)	0.9192 (47.25)	0.9046 (57.09)	0.9198 (45.42)	0.9455 (47.08)	0.276
100	50	0.2	-0.8	0.9286 (45.39)	0.9351 (51.49)	0.9226 (47.25)	0.9102 (57.11)	0.9271 (45.46)	0.9457 (47.13)	0.277

Table 5

Summary statistics for the two groups

	Estimated ICER = 15.33, NHB = 77.08					
	Control group ($n_1 = 54$)			*Intervention group* ($n_2 = 53$)		
	Mean	*S.D.*	*Skewness*	*Mean*	*S.D.*	*Skewness*
Cost (U.S.$)	2507.42	4460.44	4.93	1325.48	1785.67	3.46
Effectiveness (days with anxiety attack)	211.52	139.68	-.30	134.42	134.55	.71

Table 6

Summary of confidence intervals for the ICER and the NHB

	Methods	Confidence intervals	Interval length
ICER			
	Taylor	(-3.44 , 34.10)	37.54
	Fieller	(-1.43 , 53.86)	55.28
	Hinkley	(-1.68 , 52.44)	54.12
	Bootstrap-t	(4.50 , 49.55)	45.05
	T_1	(3.33 , 81.87)	78.54
	T_2	(-1.99 , 35.22)	37.21
	T_3	(0.35 , 38.38)	38.03
NHB			
	Normal	(25.12, 129.03)	103.92
	Bootstrap-t	(21.68, 132.33)	110.65
	BCa	(20.21, 125.52)	105.31
	T_1	(26.78, 130.87)	104.10
	T_2	(25.46, 129.39)	103.92
	T_3	(32.84, 144.42)	111.58