The Augmented Representation of the Cost-effectiveness Acceptability Curve for Economic Evaluation of Health Technology

DAJI ARAKI1,* and ISAO KAMAE1,2,3
1Division of Medical Statistics, Kobe University Graduate School of Medicine, Japan;
2Graduate School of Public Policy, The University of Tokyo, Japan;
3The Canon Institute for Global Studies, Japan

Received 3 December 2014/ Accepted 14 January 2015

Key words: Cost-effectiveness Acceptability Curve, Incremental Cost-effectiveness Ratio, Cost-effectiveness, Health Technology Assessment, Economic Evaluation, Bootstrapping, Computer Simulation

New schemes on the cost-effectiveness acceptability curve (CEAC) were developed, which can make the CEAC augmented to be more informative regarding the types of acceptance and statistical inference. Theoretical approaches have been undertaken to address two questions: 1) how the area under the curve (AUC) can be zoned by different types of acceptance displayed on the incremental cost-effectiveness plane, and 2) how the accepted dataset of incremental cost-effectiveness ratios (ICERs), which are generated by simulation runs, can be statistically associated with a threshold of ICER for acceptance. To address the first question, the AUC of a typically sigmoid-shaped CEAC was divided into three zones according to the three segmentations of the scattered plots accepted at South-east, North-east and South-west quadrants on the incremental cost-effectiveness plane. A solution for the second question was "a new CEAC of the mean" (mCEAC), which is defined by plotting a pair of the mean and its occurrence probability of ICER accepted at North-east quadrant on the incremental cost-effectiveness plane. All those schemes were graphically illustrated based on hypothetical examples using the bootstrapping simulation. Our new schemes on CEAC will provide decision makers with useful information on cost-effectiveness assessment beyond the standard presentation of CEAC.

Health Technology Assessment (HTA) has been one of the major focuses for value-based policy making in healthcare since the National Institute for Health and Care Excellence (NICE), HTA agency in England, was established in 1999. According to the Pharmacoeconomic guidelines of the NICE, it is recommended to use the cost-effectiveness acceptability curve (CEAC) as a probabilistic method which can evaluate uncertainty of the cost-effectiveness evidence obtained from the cost-effectiveness analysis (12). Canadian HTA agency, CADTH, which is another leading HTA agency in the world, also recommends the visual presentation of CEAC (4). Those recommendations by NICE and CADTH made the CEAC the most important methodology in the cost-effectiveness analysis. As the results, published articles of cost-effectiveness analysis commonly report the CEAC presentation (3, 6, 7, 9, 11, 13, 14, 15, 16, 17, 18, 19, 21).

In academic journals, the CEAC introduced in 1994 (20), has been established as a general solution to the problem of presenting uncertainty in decision modeling (2) and cost-effectiveness decision making (10).

The International Society for Pharmacoconomics and Outcomes Research (ISPOR) refers to the CEAC in the Book of Term, stating that "CEAC plots the probability that one treatment is more cost-effective than another, as a function of the threshold willingness to pay for one additional unit of efficacy. The CEAC is a graphical expression of the cost-effectiveness comparison between two treatments", as shown in Figure 1 (1). The detailed explanation in the ISPOR Book of Term is as follows1:

"Let the difference in mean costs between Treatment 2 and Treatment 1 be denoted by ΔC, and let the difference in mean efficacy (or effectiveness) between Treatment 2 and Treatment 1 be ΔE. Then the familiar incremental cost-effectiveness ratio (ICER) for Treatment 2 against Treatment 1 is ΔC/ΔE. The ICER is traditionally compared with a threshold willingness to pay for a unit of efficacy (such as QALYs saved) K, such that if the ICER is less than K then Treatment 2 should be accepted as more cost-effective than Treatment 1.

1 HEALTH CARE COST, QUALITY, AND OUTCOMES: ISPOR BOOK OF TERMS®, Copyright 2003. Reprinted by permission of the International Society for Pharmacoconomics and Outcomes Research at www.ispor.org (ISPOR). All rights reserved.

E-mail: arakidaiji@gmail.com
The simplest way to think about the CEAC is in terms of rotating the sloping line in the cost-effectiveness plane, shown Figure 2. The line has slope $K$, and the probability of positive net benefit is always the probability that the true cost effectiveness point $(\Delta E, \Delta C)$ lies below the line. When $K=0$, the line is horizontal, and the CEAC is therefore the probability that Treatment 2 is cheaper than Treatment 1. As $K$ increases toward infinity, the line rotates to become vertical and the CEAC is then the probability that Treatment 2 is more effective. At all intermediate values of $K$, the CEAC represents a balance between cost and efficacy."

Although the CEAC is well recognized and accepted as a useful tool, still some of experts warn about pitfalls and limitations of the CEAC representation. Fenwick et.al. (5) reported that the curving configuration of CEAC can be so various that it could be sometimes misleading, depending on the scattered area of the distribution of cost and effectiveness on the cost-effectiveness plane. Also, Groot KB et.al. (8) suggested potential need for improvement of the interpretation to overcome the limitations of CEAC.

In our study, therefore, we developed new schemes to answer the following questions on CEAC which has not been well addressed so far:

1) How can the area under the curve (AUC) be zoned by the types of acceptance with an occurrence probability of each type?
2) How can the distribution of ICERs virtually generated by computer simulations be statistically associated with a threshold ratio $\lambda$ (i.e., noted as $K$ in Figure 1) maximally accepted by society?

Through the answers to those questions, the conventional CEAC can be augmented to be more informative regarding the types of acceptance and statistical inference.

**METHODS**

A theoretical approach was undertaken to address two questions raised in the introduction, and computer simulations were conducted to explore how the developed methods can be visually represented in addition to the conventional graph of CEAC.

**For question 1**

Regarding the first question, three quadrants on the plane were considered: south-east (SE), north-east (NE) and south-west (SW). According to those quadrants, the AUC of a typical sigmoid-shaped CEAC can be divided into three zones, each of which vertically represents the proportion $p_i (= n_i / n(\lambda))$ from the bottom of the horizontal axis, where $n_i$ is the number of plots accepted in the quadrant $i$ ($i = SE, NE$ and SW), and $n(\lambda)$ is the total number of plots (i.e., $n(\lambda) = n_1 + n_2 + n_3$) accepted for the threshold ratio $\lambda$.

Figure 3-a illustrates that the area under the threshold line of $\lambda$ can be segmented into three components of segment A, B and C while the threshold line of $\lambda$ rotates. Accordingly, the AUC of standard CEAC can be broken down into three zones with two curves: one bordering between zone A and B, and another between zone B and C as shown in Figure 3-b.

The configuration of the three curves in Figure 3-b varies depending on the relative position and the size of eclipse of distribution in Figure 3-a. Therefore, we examined how the segmented CEAC appears to be different according to the different patterns of eclipse. The following four patterns were adopted for that purpose:

![Figure 1. A sample CEAC.](image)

![Figure 2. The cost-effectiveness plane.](image)

Figure 1 and 2 are quoted and modified from the “ISPOR Book of Terms”
Figure 3-a. Three segments and rotating threshold.
Segment A: a segment including the plots accepted for the threshold $\lambda$ at SE quadrant.
Segment B: a segment including the plots accepted for the threshold $\lambda$ at NE quadrant.
Segment C: a segment including the plots accepted for the threshold $\lambda$ at SW quadrant.
Eclipse shows the area in which all the plots are scattered with two dimensions of cost and effectiveness.

Figure 3-b. Three zones in the AUC of CEAC.
Zone A: a sub-area of the AUC of standard CEAC, which indicates the proportion of plots belonging to segment A among whole scattered plots.
Zone B: a sub-area of the AUC of standard CEAC, which indicates the proportion of plots belonging to segment B among whole scattered plots.
Zone C: a sub-area of the AUC of standard CEAC, which indicates the proportion of plots belonging to segment C among whole scattered plots.
The border between zone A and B indicates the proportion of the plots accepted at SE quadrant among all the plots.
The border between zone B and C indicates the sum of proportions of the plots accepted at SE and NE quadrants among all the plots.
Table I. Hypothetical data set of costs and QALYs.

<table>
<thead>
<tr>
<th>No. of patient</th>
<th>SOC QALYs</th>
<th>Costs ((x,1000,\text{USD/QALY}))</th>
<th>No. of patient</th>
<th>New Treatment QALYs</th>
<th>Costs ((x,1000,\text{USD/QALY}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.11</td>
<td>4</td>
<td>6</td>
<td>0.15</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>0.22</td>
<td>2</td>
<td>7</td>
<td>0.23</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>0.33</td>
<td>6.6</td>
<td>8</td>
<td>0.3</td>
<td>29</td>
</tr>
<tr>
<td>4</td>
<td>0.44</td>
<td>13</td>
<td>9</td>
<td>0.56</td>
<td>35</td>
</tr>
<tr>
<td>5</td>
<td>0.55</td>
<td>24</td>
<td>10</td>
<td>0.7</td>
<td>20</td>
</tr>
</tbody>
</table>

SOC: Standard of care

1) the proportion of SE quadrant is relatively small,
2) the proportion of SW quadrant is relatively small,
3) the proportion of SE quadrant is relatively large,
4) the proportion of SW quadrant is relatively large.

Examples suitable for those different patterns were identified and selected from a set of scattered plots, each of which can be generated by simulation run using randomized bootstrapping with 100 runs.

Bootstrapping entails drawing a large number of bootstrap samples such as 1,000 runs or more. However, less number of bootstrap samples with 100 runs are drawn in our simulation to visualize the different patterns of scattered plots as typical examples.

For question 2

A solution for the second question is to depict a new CEAC of the mean. The authors call it the mean CEAC (mCEAC), which can be constructed by plotting a pair of \(m_i\) and \(p_i\), where \(m_i\) is the mean ICER of all the ICERs accepted in the quadrant \(i\) for the threshold ratio \(\lambda\).

As there are three quadrants of NE, SW and SE, one mCEAC can be generated according to each quadrant. However, regarding the fact that the values of all the ICERs accepted in the SE quadrant are negative, and also that the ICERs at the quadrant NE and SW cannot be simply compared even if they have the same sign (i.e., positive), decision makers must be most interested in the mCEAC constituted at the NE quadrant only. Hence we pursued a mCEAC as a solution for the question 2.

Bootstrapping simulations for the CEAC zoning for question 1 and the mCEAC generation for question 2 were performed using Microsoft Office Excel 2010, commonly based on the hypothetical data set of costs and QALYs shown at Table I. The confidence interval of the mean ICER in mCEAC was constructed by t-test.

RESULTS

For question 1

Bootstrapping simulations with 100 runs resulted in a distribution of scattered plots, illustrated in Figure 4-a, on the incremental cost-effectiveness plane, and the CEAC was subsequently drawn as shown at the right in Figure 4-a. The scattered plots of Figure 4-a is composed of 100 plots, each of which shows a pair of the mean differences of cost and QALY between two treatment groups generated by each bootstrapping. Each bootstrapping was conducted by re-sampling method in five times per regimen with replacement so that the re-sampling size for each treatment regimen can be the same as the number of patients for each treatment group initially assumed as shown Table I. The acceptability curve indicated the new therapy could be accepted with the probability of 0.44 for a threshold ratio \(\lambda=50,000\,\text{USD/QALY}\).

The three zones generated from the same CEAC are shown in Figure 4-b. They can provide more information than the single curve of the standard CEAC in Figure 4-a. The border line between zone A and B indicates the proportion of the accepted plots at the SE quadrant. Since any plots at the SE quadrant exist below and to the right of the threshold line with the slope equal to \(\lambda\), the proportion of segment A is constant with probability \(p_A\), in this case, of 0.06. The vertical distance between two borders at lower (between A and B) and upper (between B and C) position implies an occurrence probability of accepted cost-effectiveness at the NE quadrant, for example, 0.32 for the threshold ratio \(\lambda\) of 50,000 US$ / QALY. The portion of the probability for acceptance at the SW quadrant, i.e., \(p_C\), was observed to be small with 0.06 for the same ratio \(\lambda\) of 50,000.
Figure 5 through 8 shows selected examples of a pair of scattered plots and zoning of the CEAC resulted from simulation runs, each case of which corresponds to the pattern 1 to 4 described in the section of Methods. Actually we found the zoning pattern was evidently different between Figure 5 and Figure 6, although the standard CEAC curves in Figure 5 and Figure 6 were quite similar. Also the standard CEAC curves in Figure 7 and Figure 8 appeared to be similar, but the zone profile of each figure was quite different.

For question 2
A resultant case by simulations with 1,000 runs of bootstrapping is shown in Table II with the arithmetic mean of the accepted values of ICERs at the NE quadrant and 95 per cent confidence limits of the mean. For this simulation, the large number of runs was set to make the bootstrapping more practical than the case for the question 1. In the case of \( \lambda = 50,000 \), as a result, the estimated mean was 29,900 US$ / QALY and 95 per cent confidence interval (28,500, 31,300). This result implies that the mean ICERs accepted at the NE quadrant is significantly much smaller than a threshold ratio assumed as the maximal limit for societal acceptance such as 50,000 US$ / QALY. The mCEAC generated by using the data set at Table II is shown in Figure 10, which is based on the scattered plots and zoning of the CEAC, as shown Figure 9.

---

For question 2
A resultant case by simulations with 1,000 runs of bootstrapping is shown in Table II with the arithmetic mean of the accepted values of ICERs at the NE quadrant and 95 per cent confidence limits of the mean. For this simulation, the large number of runs was set to make the bootstrapping more practical than the case for the question 1. In the case of \( \lambda = 50,000 \), as a result, the estimated mean was 29,900 US$ / QALY and 95 per cent confidence interval (28,500, 31,300). This result implies that the mean ICERs accepted at the NE quadrant is significantly much smaller than a threshold ratio assumed as the maximal limit for societal acceptance such as 50,000 US$ / QALY. The mCEAC generated by using the data set at Table II is shown in Figure 10, which is based on the scattered plots and zoning of the CEAC, as shown Figure 9.

---

**Figure 4-a.** Scattered plots and the standard CEAC.

**Figure 4-b.** Zoning of the CEAC in the case of Figure 4-a.
Figure 5. Scattered plots and zoning of the CEAC: Example 1.

Figure 6. Scattered plots and zoning of the CEAC: Example 2.

Figure 7. Scattered plots and zoning of the CEAC: Example 3.

Figure 8. Scattered plots and zoning of the CEAC: Example 4.
## Table II. Probability and the mean ICER with 95% CI based on Example 5.

<table>
<thead>
<tr>
<th>Threshold (x 1000 USD/QALY)</th>
<th>Probability</th>
<th>Mean (x 1000 USD/QALY)</th>
<th>95% confidence interval (lower, upper)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.087</td>
<td>6.2</td>
<td>(5.1, 7.4)</td>
</tr>
<tr>
<td>20</td>
<td>0.131</td>
<td>11.6</td>
<td>(10.5, 12.7)</td>
</tr>
<tr>
<td>30</td>
<td>0.196</td>
<td>18.3</td>
<td>(17.1, 19.8)</td>
</tr>
<tr>
<td>40</td>
<td>0.264</td>
<td>24.7</td>
<td>(23.5, 26.2)</td>
</tr>
<tr>
<td>50</td>
<td>0.360</td>
<td>29.9</td>
<td>(28.5, 31.3)</td>
</tr>
<tr>
<td>60</td>
<td>0.431</td>
<td>34.7</td>
<td>(33.1, 36.2)</td>
</tr>
<tr>
<td>70</td>
<td>0.476</td>
<td>37.9</td>
<td>(36.2, 39.5)</td>
</tr>
<tr>
<td>80</td>
<td>0.511</td>
<td>40.7</td>
<td>(39.0, 42.5)</td>
</tr>
<tr>
<td>90</td>
<td>0.543</td>
<td>43.6</td>
<td>(41.7, 45.6)</td>
</tr>
<tr>
<td>100</td>
<td>0.558</td>
<td>45.2</td>
<td>(43.2, 47.2)</td>
</tr>
</tbody>
</table>

Figure 9. Scattered plots and zoning of the CEAC: Example 5.

Figure 10. Mean CEAC and 95% confidence intervals for the ICERs based on Example 5.

\[ \lambda \]: Threshold of acceptance for incremental cost-effectiveness ratios
DISCUSSION

Our study aimed at enhancing the standard CEAC by adding more information such as segmentation by quadrants or 95% confidence intervals for the mean ICERs. As shown in Figure 4-a, conventional interpretation for the likelihood of the threshold of 50,000 US$ / QALY or less would be 44%. Our approach, however, brings more information with detailed probabilities of 6% for zone A, 32% for zone B, and 6% for zone C shown in Figure 4-b.

Such a break-down of probabilities in details is evidently advantageous for decision makers in health care, because decision makers less prefer the zone C with clinical benefit decreased, and want to avoid the risk of decreasing the clinical benefit regardless the cost. Therefore, it may be fair to say that the cost-effectiveness of a new therapy less than 50,000 US$ / QALY would be 44%, but still there is a risk of decreasing the clinical benefit of patients with a probability of 6%. It means decision makers should recognize that substantial likelihood of attaining increased benefit for patients would be 38%, not 44%. Advantageously our method enables us to make such estimation.

Contrary to the case of zone C, the distribution of plots in the zone A is the most favorable for decision makers, because the new therapy in zone A provides more benefit with less cost. Hence, decision makers are always welcome the cases in zone A, and should be careful to assess how much such dominant outcomes (i.e., more benefit and less cost) could happen by applying a new therapy. In the case of Figure 4-b, our approach can clarify that it is 6%, implying a small probability, though the standard CEAC does not indicate any information of such dominance.

Regarding the dominant outcomes, as described in the section of Results, the comparison between Figure 5 and Figure 6 illustrates the advantage of our segmentation scheme. Although the standard CEAC curves in Figure 5 and Figure 6 were quite similar, apparently decision makers prefer Figure 6 to Figure 5 by recognizing the chance of 10% for dominance in Figure 6, but almost zero in Figure 5. The zoning of CEAC can provide such information beyond the standard CEAC.

Similarly, the comparison between Figure 7 and Figure 8 gives us a lesson that the contribution of zone C not always small, and could be considerable in the case of Figure 8. In Figure 8, the crude estimate of acceptance is indicated with about 50% for the threshold of 50,000 US$ / QALY. The segmented estimate of acceptance, however, indicates about 40% at the border curve between zone B and C for the same threshold. It implies the net probability of acceptance excluding the cases with less clinical benefit must be 40%. Decision makers should know this difference of 10% to properly interpret the probability of acceptance, but they might overestimate it, regarding the crude estimate of 50% as the probability of acceptance with increased benefit. Using the zoning of CEAC enables us to avoid such misleading. Therefore, reminding us of a question, whether the zone C is substantial or not, would be another advantage of the zoning scheme.

Decision makers would be the most concerned about zone B, because there is still risk of statistical uncertainty even if they are told the probability of zone B less than the threshold of 50,000 US$ / QALY is 36% in the case of Table II. The estimated mean of ICERs in the zone B, in that example, is 29,900 US$ / QALY, which is much smaller than 50,000, and its 95% confidence interval is (28,500 , 31,300). It implies how the difference between the mean ICER of 29,900 and the threshold of 50,000 is statistically associated. It is obviously advantageous for the mCEAC to overcome the claim that the standard CEAC tells us no information of statistical significance on the ICER distribution.

Major limitation of our scheme would be no gold standard for the appropriate number of simulation runs in either zoning of standard CEAC or 95%CI representation by mCEAC. As is the case in bootstrapping, there is the same problem with Monte-Carlo simulation, which is often performed in probabilistic sensitivity analysis of economic evaluations. Since those two methods may generate different types of scattered plots, it is left for further investigation in the future to clarify the relation between our new schemes and a selection of probabilistic computing method, either bootstrapping or Monte-Carlo simulation.

Nevertheless, our current schemes, whether zoning of CEAC or drawing of mCEAC, do not necessarily depend on the selection of either bootstrapping or Monte-Carlo, because we assume that a set of scattered plots is given by some sort of method, either bootstrapping or Monte-Carlo simulation. It is commonly observed in both methods that the bigger the number of simulation runs becomes, the narrower the confidence interval does. Hence, too large number of runs makes the estimation of 95%CI practically meaningless in both methods, whether bootstrapping or Monte-Carlo simulation. With respect to the configuration of CEAC, several atypical curves are known such as a reverse sigmoid curve. Even if the shape of CEAC is different from the case that we investigated in this paper, the same principles would be applied for atypical cases although detailed exploration for different shapes of CEAC is left for the next challenge.

Another disadvantage of our scheme is no availability of computer software or application program specific to perform the zoning of standard CEAC and to generate mCEAC with the calculation of 95%CI. Without any good tool of computing, extra effort is required for researchers to perform the complex calculations. So it is
expected that a program module capable of generating zoning and mCEAC will be developed and become available in the market in the future.

As the methodology of cost-effectiveness assessment in health care has been well recognized in, not only researchers, but also medical professionals and decision makers, further improvement will be required for them to make proper interpretation of the methods. Certainly, visualizing three zones of CEAC and presenting the mCEAC will provide decision makers with more useful information in terms of their interpretation on the cost-effectiveness assessment than the depth to which the standard CEAC has ever attained.

ACKNOWLEDGEMENTS

None of the authors has any conflicts of interest or any financial ties to disclose with respect to this research.

REFERENCES


E17