Reviewer's report

Title: Using and interpreting cost-effectiveness acceptability curves: An example using data from a trial of management strategies for atrial fibrillation

Version: 1 Date: 16 January 2006

Reviewer: John William W Stevens

Reviewer's report:

General

1. This paper does not appear to provide anything new on the subject of using and interpreting cost-effectiveness acceptability curves. Its impact might be enhanced if it was based on a more extensive literature review.

2. The authors acknowledge in several places that the CEAC gives the probability that the new intervention is cost-effective relative to the alternative given the observed data for a range of values for the willingness to pay. Such a probability only has meaning from a Bayesian perspective as O'Hagan et al[2] point out, although they do acknowledge the circumstances when a frequentist interpretation is possible. The authors might wish to avoid addressing the issue of Bayesian versus frequentist statistics but it is hard to see how this is possible when interest is in making probabilistic statements about population parameters.

3. The use of the AFFIRM data to illustrate the construction, application and interpretation of the CEAC is not very illuminating because the rate-control treatment is almost the dominant strategy. There is minimal uncertainty even for an infinite willingness to pay, and much less than many decision-makers would routinely face in practice.

4. The main contribution to the literature appears to be the reporting of the results of the AFFIRM study.

--------------------------------------------------------------------------------------------------------------------------

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. Page 5:
The authors report a p-value for the mean survival gain, which is assumed to be a two-sided p-value. However, a frequentist interpretation of a CEAC is usually a one sided test of the null hypothesis that the net benefit is not acceptable against the alternative hypothesis that the net benefit is acceptable. This distinction is not made clear by the authors.

2. Page 5:
The survival effect size in AFFIRM is small, about one month, and the uncertainty, which is relatively small, suggests that the true treatment effect could be a decrease of one month or an increase of 3 months. Although in most cases it is important to look at the joint distribution of costs and effects when making inferences about cost effectiveness, I am not persuaded that inferences are affected by this in the case of the AFFIRM study; there is no trade-off to be had between costs and effects in the north-east and south-west quadrants and very little statistics is required.
3. Page 8:
The authors should make it clear that the bootstrap is generating the sampling distribution of the joint mean cost and efficacy.

4. Page 9:
There is nothing inherently wrongly with using the joint distribution to make inferences about marginal effects if these are of interest.

5. Page 9:
The problems associated with calculating a confidence interval for an ICER are not resolved, as the authors suggest, through using Fieller’s theorem or the non parametric bootstrap when the denominator is close to zero[1,2].

6. Page 11:
As the authors indicate, the CEAC is a probabilistic statement about population parameters. When analysing a clinical trial the frequentist approach is typically to use the non-parametric bootstrap, whereas when analysing an economic model the approach is to use Monte Carlo simulation to propagate uncertainty about the outputs through uncertainty about the inputs. This distinction is unclear from the authors’ description of the derivation of the CEAC.

7. Page 12:
It is unclear why the authors choose to mention that the CEAC does not represent a cumulative distribution function.

8. Page 12:
100% of the cost-effect pairs do not fall to the south and east of the line when I was zero, they fall to the south of the line.

REFERENCES


What next?: Accept after minor essential revisions

Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Acceptable

Statistical review: Yes
Declaration of competing interests:

I hold shares in AstraZeneca and GlaxoSmithKline, which develop treatments for atrial fibrillation