Author’s response to reviews

Title: The Probability of Cost-Effectiveness

Authors:

Anthony O'Hagan (a.ohagan@sheffield.ac.uk)
John W Stevens (John.W.Stevens@astrazeneca.com)

Version: 2 Date: 26 Feb 2002

PDF covering letter
Response to referees

Paper Title: The probability of cost-effectiveness.

Authors: Anthony O’Hagan, John W Stevens

General comments

Of the two reviewers, Professor Heitjan expresses complete satisfaction with the paper and advises acceptance without revision. Dr Willan, who might reasonably be said to have a competing interest as the author of the paper that we criticise, makes a number of comments and demands that we make revisions in respect of every one. We suggested Dr Willan as a referee, and are pleased to see that the journal has followed this suggestion. Some of his contributions have merit, and we feel that in revising our paper to reflect those comments it has been substantially improved. We hope that he will agree with us.

Although Professor Heitjan had no concerns with the previous version, we suggest that he should see Dr Willan’s comments and our responses because we would value any further input from him.

Dr Willan’s comments

Dr Willan makes the following General Comments:

The authors take issue with a definition for the probability of cost-effectiveness proposed by Willan, and propose one of their own. The main difference reflects the fact that the authors take a Bayesian approach for which probability statements about population parameters are permitted, while the definition they criticise is based on a frequentist approach for which probability statements are confined to events.

This is inaccurate in a number of ways. We do not propose any new definition for the probability of cost-effectiveness, only give our support to the definition in terms of the Cost-Effectiveness Acceptability Curve (CEAC) that Willan criticised. We did not object to his definition per se, but disagreed with his assertion that it was a more appropriate definition than that entailed in the use of the CEAC. It is still our firm opinion that the definition of the probability of cost-effectiveness as a measure of certainty, based on the available data, that one treatment has a higher population mean net benefit than another, is the one most useful to decision makers. We therefore stand firmly by our original position.

It is not a matter of Bayesian versus non-Bayesian approaches. We took some space in our paper to address the distinction between the two approaches because we felt that some of his assertions were inaccurate or misleading. Whilst we are both entirely committed to the Bayesian approach, we did not in this paper intend to take any partisan positions.
Of course, what seems to one of us as incontrovertible and non-partisan may be perceived quite differently by the other. This is one reason why we value Dr Willan’s input. We have tried in the new version to be even less controversial.

Dr Willan made six Specific Comments.

1. The authors claim, with no substantiation, that knowing the probability that a patient given treatment 2 has a higher net benefit than a patient given treatment 1 is of “little or no relevance to a policy maker”.

2. The authors claim that health care providers are only interested in the “aggregate of patients, rather than the individual patient”, and therefore would not be interested in knowing “the probability that treatment 2 will be more cost-effective than treatment 1 for the next patient”. I don’t believe this is true. The health care providers I deal with are most definitely interested in knowing how likely the next patient will benefit from potential interventions. Furthermore, it is not just the probability for the next patient, it is also the probability for all future patients, providing the aggregate prospective [sic] as well. Most population parameters estimated in clinical trials apply to individual patients (e.g. mean difference in cost, mean difference in effectiveness), but they are still very much of interest to the consumers of the results of [the] clinical trial since they provide estimates of the clinical significance of the difference between treatments 1 and 2.

It is completely standard in health economics that the primary issue for a health care provider is the expected net benefit comparison. What proportion of patients would experience a higher net benefit on one treatment than the other may be of interest, but it is very much of secondary importance. It may be that, for instance, 75% of patients would experience a higher net benefit under treatment 1 and yet that the expected net benefit under treatment 2 is higher. The latter fact is of prime consideration to the health care provider. It is the total good of the patients in their care that matters. The 75% figure would certainly be of some interest, and we thank Dr Willan for taking us to task on under-valuing it. We have amended our paper accordingly. However, we stand firmly by our position that Dr Willan’s probability of cost-effectiveness is of ancillary interest, and we continue to take issue with his paper on this account.

We also accept that we have used the ideas of aggregate versus individual loosely in this respect, and have tried to be more careful in the revised version.

3. The authors state, “Willan’s dislike of the fact that Q(K) depends on the data seems to be based on failing to appreciate that it is an inference, not a parameter.” The authors are mistaken. Willan makes it abundantly clear that his concern about Q(K) being labelled “the probability of cost-effectiveness” is that it is an inference. Quoting their quote from Willan’s paper, “Attaching the label ‘the probability that the intervention is cost-effective’ to this quantity could mislead policy makers into thinking that treatment is highly beneficial compared to standard. What, in fact, is high is our confidence that the INB, however small, is not zero.” And quoting
from Willan’s abstract, “These definitions are, in fact, expressions of the
certainty with which the current evidence would lead us to believe that the
treatment under consideration is cost-effective.”

4. The statement that, “This seems to indicate a low opinion of policy
makers” is not substantiated. Recognizing the fact that people can be
confused by confusing labels does not confer a low opinion.

Willan does not clearly distinguish between inferences and population parameters,
and we sought to clarify that distinction. It may have been clear in his mind but it did
not emerge clearly in our reading of the paper. We defer on this, however. Perhaps
our reading was affected by our overall concern about this paper.

However, since Dr Willan reiterated the passage that we quoted, we feel we should
point out that his definition of ‘the probability that the intervention is cost-effective’
no more implies that treatment is highly beneficial compared to standard than does
our definition. If his (estimated) probability of cost-effectiveness turned out to be
0.99 (the figure suggested by Willan in the sentence preceding the quotation), then
what, in fact, is high is (an estimate of) the proportion of patients in the population
whose INB, however small, is not zero. If policy makers were in danger of
misreading the definition in terms of the CEAC this way, then they are in equal
danger of misreading his definition in exactly the same way.

Our paper did attempt to clarify the issues by setting out the distinction between a
probability as an inference (in Bayesian statistics) and an unknown population
proportion (which is also called a probability in frequentist inference). We hope that
the new version is even more successful in this respect.

In an attempt to resolve our differences with Dr Willan, we suggest that the language
used by both Bayesian and frequentist analyses should be more precise. This is now
proposed in the revised paper.

5. I agree with the authors when they state, “Our certainty, based on available
evidence, that treatment 2 is more cost-effective than [sic] treatment 1 is
properly expressed by a probability” whether it be a p-value or a Bayesian
posterior probability. That does not mean that consumers of the results of
clinical trials will not be confused by the label “the probability of cost-
effective”.

We would like to remind Dr Willan that a p-value is not the probability that the null
hypothesis is true, and must never ever be misread as such.

6. The authors make a good point when they take issue with the estimator
proposed by Willan. Because of the likely positive correlation between a
patient’s observed net benefit on treatment 1 and his or her net benefit on
treatment 2, the proposed estimator is biased toward 0.5. Therefore, if in
truth net benefit on treatment 2 is larger than on treatment 1, Willan’s
estimator provides a lower bound for the estimate. It does not mean,
according to the authors that, “it is completely impossible to learn about
p(K).” Nor does it mean, “his proposed estimate of this probability is
logically flawed.” Using the data from a clinical trial, one could easily provide a sensitivity analysis over a range of correlations.

Our statement that, “it is completely impossible to learn about $p(K)$” has been amended slightly. As Dr Willan points out, it is possible to gain some information about $p(K)$, but the fact remains that no matter how much data we obtain we can never learn more than that it lies in a range.

By not mentioning the matter of correlation, Willan’s paper is most definitely wrong, and his proposed estimate is most definitely logically flawed. We cannot see how he can dismiss our claim. Even if it were the case that a different procedure (that was not mentioned in the paper) might not be so flawed, this would not in any way make the original estimator logically sound.

Dr Willan’s new suggestion that the original estimator provides a lower bound goes too far and does not resolve the underlying flaw. We said, “Indeed one might imagine in practice there would be quite strong correlations, on the basis that a patient who responds well to one treatment might respond relatively well to the other, and similarly for costs.” We were careful not to suggest that this implied a positive correlation between net benefits. There is nearly always a positive correlation between costs and effectiveness, and even with all these correlations positive it is possible for net benefits to be negatively correlated. Anyway, it would not in our opinion be proper to assume any positive correlation in a context where there is not, and cannot be, any evidence about it in the data.

Dr Willan speculates on sensitivity analysis, but we suspect the details may be complex. Whatever the merits of the proposals he now makes, and whether or not they might after due and careful mathematical development lead to a fresh paper that proposes an estimator that actually has the properties claimed for it, the fact is that none of this was mentioned in his original paper. His method in the paper was wrong, and we are absolutely within our rights to point that out. Indeed, we see it as our duty to those people, involved in the assessment of cost-effectiveness, whom this paper was intended to influence.

Tony O’Hagan
John Stevens