Reviewer's report

Title: A Systematic Review on Methods of Health Research Economic Impact Evaluation

Version: 3 Date: 21 November 2009

Reviewer: Steven Wooding

Reviewer's report:

1) From Discussion: I see every reason why we should be careful in interpreting studies on the benefits of research but I see no reason why we should be any more sceptical of results that suggest investment in research isn't worthwhile, than in ones that suggest it is. Surely if we are to be scientific about this we shouldn't be expecting an answer that is one way or the other. "Care should be taken in interpreting research results that undermine the economic benefits of research, and such results should not always be interpreted as a waste of resources."

2) The article makes a distinction between macro economic studies and case study research - but it gets the distinction wrong on a number of occasions - specifically the only work explicitly referred to as case study research is the 'Evaluation Forum' report, this was not case study research - both the American Exceptional returns work and the UK Evaluation Forum work looked to estimate the return on investment of cardio-vascular research - both calculated the expenditure on research and estimated the benefits. The key difference was that the Exceptional returns work took no account of lags and calculated benefit from overall drops in population mortality, whereas the Evaluation forum work estimated the average lag time for benefit and identified the healthcare benefits from individual treatments.

The article then goes on to talk about frameworks - so doesn't talk about case study research specifically again research. Although most of the frameworks can be used to structure case studies (or survey work) - particularly the Payback Framework. The article doesn't talk about the comparison of macroeconomic (estimates over rate of return, but doesn't tell you how it happens) and case studies (tells you a lot of detail about what happened and provides ideas about how to improve the system; but doesn't allow estimates of overall rate of return).

I would say the structure should be a) macroeconomic studies that are primarily quantitative and carried out across sectors or science as a whole and b) more qualitative case study or survey work which often employs a logic model style framework and classification of benefits, the most common frameworks for which
are discussed.

The article suggests it will make this distinction in the abstract: "There are two basic methods for studying such benefits: macroeconomic evaluation of the relationship between funds spent on research and economic benefits, and case studies which examine a single research or program."

3) Section 1 discusses the American and Australian exceptional returns and related studies and suggests that the major limitation of such studies is confounding factors and attribution. I think there is a danger of confusion between those particular studies and this type of study in general. I would say the major weaknesses were, for the American and first Australian study, neglecting the lag between research and benefit; for the second Australian study huge assumptions about the future health benefit produced by research. If these limitations were addressed then the attribution and confounding factors would be the key challenge.

4) The argument about including implementation costs is also misses the key point that in general it is essential to include the costs of implementation when comparing to the benefits provided - it doesn't only matter in cases where the benefits are likely to be low - it matters for all improvements when comparing them to other interventions or ways of spending public money. The case given is the extreme one, but it is not presented as such: "Some researchers believe that the expense of implementing research results should be considered in evaluation of the economic benefits of such research because high expense is of little benefit for patients who are near death; research on novel care techniques has fewer benefits than research on preventative methods; and theoretical estimates have shown that, for specific gender and age groups, the benefits of prolonging life are less than curative expenditures [25]."

5) I'm confused about how the Conclusions link to the data presented in the paper:
   A) the paper suggests that methods need to be sensitive to objectives and context; the conclusions call for standard methods
   B) the paper makes no mention of the benefits of being able to compare large numbers of studies or cases and doesn't talk about the challenges of doing this when methods are non-standard; but yet the conclusion calls for standard methods
   C) the paper does a reasonable job of laying out the types of approaches used to assess the benefits of healthcare research, most of which have been developed
by particular researchers on institutions; and then suggests that international organisations should use a set of methods (CHOICE) that have not been covered or discussed in the article (or as suggested in the initial abstract "We suggest that methodological tools be designed at the international level to assess the economic impact of healthcare research in specific settings") - why recommend the development of new tools rather than use the CAHS or the Payback framework?

- Minor Essential Revisions

1) Research Impact Framework - who was it introduced by and for what use?

2) I'm don't think time scale terms such as 'short term' in the following are useful - it would be better to give a time scale eg 5-10 years - I would not count 5-10 years as short term, but in my experience such studies are normally carried out over at least this time scale: "Different methods have been proposed to identify the impact of healthcare research on human health for the purpose of differentiating it from other factors. One proposal is the study of short-term effects of specific efforts, such as cardiovascular research."

3) "An attribution of 50% is considered acceptable in some studies, and a sensitivity analysis of 30-70% has been used to obtain a higher degree of certainty [11]." It should be stressed that these numbers are guesses or at best guesses estimates and the use of 30-70% does not increase the degree of certainty - it more accurately reflects the degree of uncertainty in the initial guesses estimates.

4) "One study used ‘additional earnings of cancer survivors’ to this end [27]." I have not checked the study referenced but I would suspect that this measure is a 'productivity approach' measure and hence should be included in this category.

5) "However, after examining various studies, the ‘UK Evaluation Forum' report in 2008 suggested 10-25 years (average: 17 years) for cardiovascular research and 9-14 years (average: 12 years) for mental health research [12]." The report tried to estimate the average time from research to health impact - this is not quite the same as recommending the timeframe of evaluation, ie it is the time frame to see 50% of the impact on health policy which may or may not be appropriate as the time window for evaluation.

- Discretionary Revisions

1) I'm not clear why the report "Medical Research - What's it worth?" is referred to as the 'UK Evaluation Forum' study whereas the Exception Returns work is referred to by it's title.
2) I would argue that sections on economic benefits are the least convincing because often there is an absence of good data - I'm not sure that is the same as methodologically weak - i.e., the methodology is irrelevant if there is no data. "In the studied reports, the sections on evaluation the economic impacts of healthcare research are often the weakest methodologically,"

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

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I work for an organisation that carries out significant work in the area of research evaluation hence the overall fate of the field may affect the fortunes of the organisation.