Reviewer’s report

Title: The cost-effectiveness of tracking newborns with bilateral hearing impairment in Bavaria: A decision-analytic model

Version: 2 Date: 7 June 2012

Reviewer: Andreas Gerber

Reviewer’s report:

The cost-effectiveness of tracking newborns with bilateral hearing impairments in Bavaria.

Dear authors,

thanks for your health economic evaluation of a screening and tracking program for newborns to detect bilateral hearing impairment.

Overall I must say that I do not see a basis to perform a health economic evaluation of the problem you address. The data as you know are very weak, see Wolff et al. 2010 and IQWiG 2007 – with reference to these publications I would like to make you aware that they are the same product which I cannot gather from the way you quote them. Wolff p. 131 right column: “study quality was generally poor …. Only one treatment study showed “minor deficiencies”.” Therefore, we cannot currently say anything about the effect of early vs. late treatment, henceforth nothing on the effect that a tracking system would have on the reduction of impairments children with bilateral hearing impairments would suffer from. Further, as you can see from p. 134 left column the authors conclude: “… this can only be regarded as indication that the expressive and receptive language abilities …. are better in children treated earlier.” In the context of health economic evaluation of a screening program I would like to also mention the concept of linked evidence meaning that a screening in itself that is not linked to an effect in treatment would not be considered a patient-relevant outcome according to Social Code Book in Germany in particular and according to current standards of evidence-based medicine in general, see publications by Holger Schuenemann from Hamilton University on the concept of linked evidence).

Therefore, from my personal point of view and from the point of view of evidence-based medicine I think it is too early to perform health economic evaluations without the basis of solid evidence. This is a common problem of screening interventions, but we should be aware of it and not make conclusions on an invalid data basis. I think we cannot publish papers any more that do health economic analyses without knowing anything about the long-term outcomes, as you yourself point out in p. 7 of your manuscript.

Aside from this basis comment I have some more general comments. This paper seems to be drafted with a very specific audience in mind, making it difficult to
comprehend. Substantial knowledge of Germany seems to be necessary to fully appreciate. It is, for instance, never made clear what the Joint Federal Commission is and how it is relevant.

Furthermore, the presentation is a bit problematic as it is difficult to follow, in particular when the reader does not have any background in hearing impairment screening. It is not clear why several test stages are needed if the AABR test as a 100% test accuracy (p. 10). From the 9 studies on reliability and validity of methods that were included in the IQWiG 2007 report S-OAE and A-ABR not sufficiently evaluated. So please explain where you get the data from.

Specific comments:

The authors might want to describe the pilot, the number of included children from Upper Palatine etc. in more detail for the non-German native speakers that would want to read this article. A short description of the timeline etc. would do. As I know that reviewers ask for additions, but never suggest where to cut I would suggest to shorten the section on the literature review of other health economic evaluations and report them in a separate paper or just refer to some of them that are specifically on tracking in the discussion.

The authors use the term “failing” a test to indicate the test was positive (see Figure 1 footnote). In the abstract, however, they speak of “positive test results” and do not use “failing”. A positive test could be false or a true positive. I miss a discussion whether the higher detection rate with tracking also affects the ratio of true to false negative and how this affects cost-effectiveness.

It is not clear to me how the confidence intervals for the cost estimates in Table 2 were derived. In particular as the EBM system seems to be a fixed fee aside from some minor regional differences throughout Germany.

Table 4: Why would you need an epidemiologist if the tracking program became standard? Therefore I do not understand why you should include these costs?

P 13: What are the “better outcomes” of tracking?

In the discussion of the review of existing models (p. 11) it is not discussed why the estimate differ so dramatically, e.g. 25,813 GBP of a child detected vs. 5,113USD. Furthermore, currency abbreviations are not explained.

The authors report only ICERs for their analysis and not ACERs. This would not only be insightful but also helps to make it comparable with the estimates reported in the review of existing models.

The authors acknowledge four major limitations. In particular the explanation of the first limitation mentioned (the conditional probability of failing) has to be expanded. Furthermore, the discussion of the limitations should also try assessing how these affect the cost-effectiveness, i.e. on what direction are the ICERs affected.

A further major limitation I see is in the set-up of the costs and probabilities for
the model. If I understand it correctly, the probabilities for a given stage are a weighted average of different kinds of test with different corresponding failure rates. This mixture is correctly reflected when calculating the costs by ensuring the costs at each stage reflect this weighted average of the corresponding tests. This means the results are only valid if this weighted average is constant. For example, if at the 2nd stage the share of (cheaper) OAE tests increases the probabilities used are not valid anymore and it is unclear in what direction the effectiveness changes.

On p. 15: I do not understand the term “methodological uncertainty” in this context, please expand.

Some editing could increase clarity tremendously, e.g. on p. 7: “the benefits […] are open”.