Reviewer's report

Title: S-carboxymethylcysteine in the treatment of glue ear: quantitative systematic review

Authors:
Robert A Moore (andrew.moore@pru.ox.ac.uk)
Dermot Commins (dcommins@freeuk.net)
Grant Bates (grant.bates@surgery.oxford.ac.uk)
Ceri J Phillips (C.J.Phillips@swansea.ac.uk)

Version: 6 Date: 12 Sep 2001

Reviewer: Dr Matthias Egger

Level of interest: A paper of considerable general medical or scientific interest

Advice on publication: Unable to decide on acceptance or rejection until I see revised version

This is an interesting review which addresses the important question whether the mucolytic S-carboxymethylcysteine is beneficial in glue ear and prevents the insertion of grommets in children. The paper could be improved in a number of respects:

1. The quality of reporting in general is inadequate. The authors should consult the checklist of the QUORUM statement (Moher et al, Lancet 1999) and consider the inclusion of a flow diagram showing the numbers of RCTs identified, included, and excluded.

2. Many of the statements made in the introduction are not referenced ('42% of three year olds may begin an episode of glue ear over the next twelve months', 'glue ear can result in a hearing loss with an average of 20 decibels etc'). It would be helpful to cite the relevant references and thus underpin these statements.
3. It is unclear how exactly improvement was defined for this analysis. For example, in Table 1 the study by Taylor and Dareshani is described as having four outcomes including hearing levels, tuning forks, audiology but it is unclear which one was used in the systematic review and meta-analysis.

4. The assessment of the methodological quality of studies using quality scales and summary scores is problematic (see Juni et al, JAMA 1999). It is preferable to report for each study whether important dimensions of methodological quality, such as concealment of allocation and blinding of outcome assessment, were met. The Jadad score, which was used by the authors, is particularly problematic because it does not assess concealment of allocation, the dimension which has empirically been shown to be of particular importance (see for example Schulz et al, JAMA 1995). Indeed, the Jadad score considers the use of an open random number equivalent to concealed randomisation using a telephone or computer system. The second score used by the authors is not described in any detail (Smith et al, erroneously referenced as 'in press' in Table 1) but also appears to ignore concealment of allocation. The problems with scales are in fact nicely illustrated by the discordant results that were obtained with the two scales.

5. The stratified analysis by quality score (Table 2) is therefore potentially misleading - it would be much more informative to see what happens when trials are stratified according to key domains of methodological quality, including blinding and concealment of allocation, rather than scores.

6. The description of the statistical methods is incomplete. What type of fixed effects model was used? Were number needed to treat (NNTs) calculated for each study and then statistically combined (how?) or was the benefit ratio, odds ratio or risk (benefit) difference used for meta analysis and then converted into the currency of numbers needed to treat? If the latter is the case, what baseline rate of improvement was assumed? Also, the authors state that numbers needed to harm were calculated but they do not actually report these numbers. Finally, it is unclear how the authors dealt with ears and children: could the same child contribute both ears to the analysis?

7. It is unfortunate that no forest plot is shown and that the results of the individual studies (NNTs or relative benefits) are not given for each study. The authors state that the combined NNT was 5.5 with 95% confidence intervals ranging from 3.8 to 9.8. I am not convinced that this is a useful representation of the data (see also Smeeth et al, BMJ 1999). The NNT will of course depend on the rate of spontaneous
improvement among controls, which showed considerable variation between the studies included in the review (as clearly shown in the authors' L'Abbe plot). The NNTs from individual studies are (in my calculation):

Taylor & Dareshani 3.6  
Ramsden et al 7.7  
Spigno & Teatini 6.7  
Hughes 9.1  
Porru infinitely large (no benefit)  
Bonci & Bozzi 2.2  
Commins et al 11  

8. It would be better to give a combined relative benefit and then to discuss what this means in terms of NNT for different, plausible situations. It would also be interesting to explore why the spontaneous rate of resolution differed between studies. The authors are used to the context of acute pain where effects are usually large and variations in baseline risk small, but this is not the case here.

9. The analysis of the Porru trial is based on 11 children in the active group but only 3 in the control group. What happened? Should this trial have been excluded?

10. Is the Commins et al study published or unpublished?

In summary:

. The conclusions drawn are not fully supported by the data shown. In particular the methodological quality of component studies is not adequately addressed.

. The methods are not described in sufficient detail to allow replication.

. The manuscript does not adhere to the relevant reporting standards (QUORUM statement)

. The writing is fine.

**Competing interests:**

None declared.