Grommets in otitis media with effusion: an individual patient data meta-analysis

M M Rovers, N Black, G G Browning, R Maw, G A Zielhuis, M P Haggard

Aims: To identify subgroups of children with otitis media with effusion (OME) that might benefit more than others from treatment with ventilation tubes.

Methods: An individual patient data (IPD) meta-analysis on seven randomised controlled trials (n = 1234 children in all), focusing on interactions between treatment and baseline characteristics—hearing level (HL), history of acute otitis media, common colds, attending day-care, gender, age, socioeconomic status, siblings, season, passive smoking, and history of breast feeding. Outcome measures that could be studied were mean time spent with effusion (n = 557), mean hearing levels (n = 557 in studies that randomised children, and n = 180 in studies that randomised ears), and language development (n = 381).

Results: In the trials that treated both ears the only significant interaction was between day-care and surgery, occurring where mean hearing level was the outcome measure. None of the other baseline variables showed an interaction effect with treatment that would justify subgrouping. In the trials that treated only one ear, the baseline hearing level showed a significant but not pervasive interaction with treatment—that is, only with a cut-off of 25 dB HL.

Conclusions: The effects of conventional ventilation tubes in children studied so far are small and limited in duration. Observation (watchful waiting) therefore seems to be an adequate management strategy for most children with OME. Ventilation tubes might be used in young children that grow up in an environment with a high infection load (for example, children attending day-care), or in older children with a hearing level of 25 dB HL or greater in both ears persisting for at least 12 weeks.
the search strategy can be found on the ADC website. To be selected for the IPD meta-analysis, trials had to be randomised to a high standard. Trials had to include children aged 0–12 years of age with tympanometrically and/or otoscopically confirmed persistent bilateral OME, and the comparison had to be between short term ventilation tubes (VT) and watchful waiting (WW). The primary investigators of all selected trials were asked for the raw data of their trials. Trials that treated one ear and used the contralateral ear in the comparison instead of a concurrent control group were analysed separately.347

**Data collection and endpoints**

The obtained data were thoroughly checked18 for consistency, plausibility, and integrity of randomisation and follow up. Any queries were resolved by the responsible trial investigator or statistician.

Primary outcome measures were mean time spent with effusion (measured by tympanometry), hearing (measured by pure-tone audiometry or age related hearing assessment), and language development (measured by Reynell test*). Calculation of the mean time spent with effusion was based on an interpolation from type B tympanograms during the follow up visits. If a child had OME at two successive measurements, the days between these measurements were counted as days with effusion. If a child had OME at the first measurement but not at the second (or vice versa), only half of the days were counted as days with effusion. Hearing level over 500, 1000, 2000, and 4000 Hz) measured where possible by air conduction, pure tone audiometry. In trials that randomised children rather than ears the binaural average was taken. Reynell language scores were expressed as standardised z scores. As the language development was measured at different follow up times in two of the trials .

* A standardised and validated test measuring the comprehensive language development.
included, we had to assume that the mean scores at 6 and 9 months follow up could be aggregated, as well as the means at 12 and 18 months follow up. The trials that studied behaviour and quality of life used differing outcome measures, so pooling was not possible.

Effect modifiers (subgroups) that could be included in the IPD were: hearing level at baseline (n = 744), history of acute otitis media (yes/no; n = 616), upper respiratory infections (yes/no; n = 619), attending day-care (yes/no; n = 625), gender (boys/girls; n = 651), age (n = 651), socioeconomic status (educational level of mother; n = 632), siblings (yes/no; n = 634), season (winter, spring, summer, autumn; n = 650), history of breast feeding (yes/no; 464), and parental smoking (yes/no; 466). All these modifiers were measured prior to randomisation.

Analyses and statistics

In the first instance, all analyses were performed as randomised (so called “intention-to-treat” principle). Differences in mean time spent with effusion, mean hearing loss, and language development between both treatment groups were tested with the Student’s t test for independent groups.

Second, a prognostic model was made to study predictors of poor outcome, which was defined as a score worse than the median hearing loss, time with effusion, and language development. Subsequently, the individual predictors were used to study possible effect modification—that is, the effects of ventilation tubes within subgroups of patients. Fixed effect regression analyses were performed with treatment group, effect modifier (subgroups) that could be included in the IPD were: hearing level at baseline (n = 744), history of acute otitis media (yes/no; n = 616), upper respiratory infections (yes/no; n = 619), attending day-care (yes/no; n = 625), gender (boys/girls; n = 651), age (n = 651), socioeconomic status (educational level of mother; n = 632), siblings (yes/no; n = 634), season (winter, spring, summer, autumn; n = 650), history of breast feeding (yes/no; 464), and parental smoking (yes/no; 466). All these modifiers were measured prior to randomisation.

Analyses and statistics

In the first instance, all analyses were performed as randomised (so called “intention-to-treat” principle). Differences in mean time spent with effusion, mean hearing loss, and language development between both treatment groups were tested with the Student’s t test for independent groups.

Second, a prognostic model was made to study predictors of poor outcome, which was defined as a score worse than the median hearing loss, time with effusion, and language development. Subsequently, the individual predictors were used to study possible effect modification—that is, the effects of ventilation tubes within subgroups of patients. Fixed effect regression analyses were performed with treatment group, effect modifier (subgroups) that could be included in the IPD were: hearing level at baseline (n = 744), history of acute otitis media (yes/no; n = 616), upper respiratory infections (yes/no; n = 619), attending day-care (yes/no; n = 625), gender (boys/girls; n = 651), age (n = 651), socioeconomic status (educational level of mother; n = 632), siblings (yes/no; n = 634), season (winter, spring, summer, autumn; n = 650), history of breast feeding (yes/no; 464), and parental smoking (yes/no; 466). All these modifiers were measured prior to randomisation.

Analyses and statistics

In the first instance, all analyses were performed as randomised (so called “intention-to-treat” principle). Differences in mean time spent with effusion, mean hearing loss, and language development between both treatment groups were tested with the Student’s t test for independent groups.

Second, a prognostic model was made to study predictors of poor outcome, which was defined as a score worse than the median hearing loss, time with effusion, and language development. Subsequently, the individual predictors were used to study possible effect modification—that is, the effects of ventilation tubes within subgroups of patients. Fixed effect regression analyses were performed with treatment group, effect modifier (subgroups) that could be included in the IPD were: hearing level at baseline (n = 744), history of acute otitis media (yes/no; n = 616), upper respiratory infections (yes/no; n = 619), attending day-care (yes/no; n = 625), gender (boys/girls; n = 651), age (n = 651), socioeconomic status (educational level of mother; n = 632), siblings (yes/no; n = 634), season (winter, spring, summer, autumn; n = 650), history of breast feeding (yes/no; 464), and parental smoking (yes/no; 466). All these modifiers were measured prior to randomisation.
Children with more than one risk factor appeared to benefit slightly more from treatment with ventilation tubes, but the accumulation was only weak, like most of the individual risk factors. In the efficacy (as treated) analysis, large significant effects on hearing level were found between children with functioning ventilation tubes and children without or with non-functioning tubes. The mean hearing level in children treated with functioning tubes was about 6 dB HL better compared to children with non-functioning tubes, both after 6 and 12 months follow up (p = 0.001) However, also in this analysis no significant interaction effects indicating relevant subgroups were found.

**Trials that randomised ears**

In the three trials that only treated one ear and used the contralateral ear as the comparison, only baseline hearing level, age, and gender could be studied as possible indicators, and mean hearing level was the only measured outcome. Other variables and/or outcomes were not measured, or the data were so heterogeneous that they could not be pooled. Furthermore, only cases in which a ventilation tube was inserted in one ear and in whom the contralateral ear was used as the control ear (n = 160) were included in the analysis—that is, children who also had undergone adenoidectomy were excluded in the analysis. The studies were also on older children that were almost all at school. In these trials ventilation tubes did appear to be differentially effective in the ears with a worse baseline hearing level. However, if the hearing level at baseline was dichotomised at various cut-off values, only a cut-off of 25 dB HL showed an effect. After 6 months follow up, ears treated with ventilation tubes and a baseline hearing loss of 25 dB HL or greater improved 10 dB HL more than ears with a similar baseline hearing loss but which were not treated with ventilation tubes. The ears of children treated with ventilation tubes and a baseline hearing loss smaller than 25 dB HL improved only 3 dB HL more than the control ears (p = 0.28 for interaction) (fig 3B).

**DISCUSSION**

The pooled results of meta-analysable studies confirm that treatment with short term ventilation tubes produces limited hearing improvement of only short duration—that is, only as long as the tubes are in situ and patent. Children aged 3 years or younger attending day-care and children aged 4 years or older with a hearing level of 25 dB HL or greater in both ears persisting for at least 12 weeks might benefit more from treatment with such ventilation tubes. Baseline hearing level, however, did not emerge as a clear selection criterion in the way suggested by current guidelines and by prevailing professional belief—that is, average hearing level at baseline did not obviously modify the effect estimate.

Interaction results for baseline hearing level with treatment differed between the trials that randomised children and trials that randomised ears. This difference is probably due to four factors. The first is the control for individual variance achieved by analysing treatment effects within subjects, reducing measurement error. The second is that there was no switching in the trials that randomised ears, whereas in the trials that randomised children, 11–85% of the children in the WW group received ventilation tubes during follow up. The third is the distribution of the mean hearing loss at baseline. In the trials that treated only one ear, the range of the mean hearing loss at baseline was 7.5–47.5 dB HL, whereas the range in the trials that treated both ears was 17.5–60 dB HL. Consequently, the power of the subgroup analysis with a cut-off around 25 dB HL was larger in the trials that treated one ear. Furthermore, the results suggest that children with a more definite hearing loss have been included in the more recently performed randomised trials—that is, the trials that randomised children. This follows from the trend in guidelines that suggest inserting ventilation tubes only in children with persistent OME and a hearing loss $\geq$ 20 dB HL. However, even for children with marked hearing loss at baseline, treatment with ventilation tubes was only effective in the short term—that is, as long as the tubes are in situ and patent. It has to be considered whether the absence

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Mean binaural hearing levels during follow up in trials (k = 4) that randomised children (pooled effect estimate)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment group</td>
<td>No.</td>
</tr>
<tr>
<td>0 months follow up</td>
<td>VT</td>
</tr>
<tr>
<td>WW</td>
<td>278</td>
</tr>
<tr>
<td>6 months follow up</td>
<td>VT</td>
</tr>
<tr>
<td>WW</td>
<td>189</td>
</tr>
<tr>
<td>12 months follow up</td>
<td>VT</td>
</tr>
<tr>
<td>WW</td>
<td>181</td>
</tr>
<tr>
<td>18 months follow up</td>
<td>VT</td>
</tr>
<tr>
<td>WW</td>
<td>135</td>
</tr>
</tbody>
</table>

VT, ventilation tubes; WW, watchful waiting; HL, hearing level; SE, standard error.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Mean language development (standardised z score) during follow up (k = 3 trials)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment group</td>
<td>No.</td>
</tr>
<tr>
<td>6/9 months follow up</td>
<td>VT</td>
</tr>
<tr>
<td>WW</td>
<td>188</td>
</tr>
<tr>
<td>12/18 months follow up</td>
<td>VT</td>
</tr>
<tr>
<td>WW</td>
<td>155</td>
</tr>
</tbody>
</table>

VT, ventilation tubes; WW, watchful waiting; SE, standard error.
of interaction between treatment and baseline hearing is due
to insensitive outcome measurement. Hearing level, however,
is unusually precise as an outcome measure.

The interaction effect that distinguishes children according
to day-care attendance might be explained by the infection
load in day-care. The infection load in day-care may result in
an increased risk of new episodes with effusion.

Consequently, children attending day-care and who are not
-treated with ventilation tubes might suffer from more
-frequent or longer hearing losses over a longer period.
However, if this hypothesis was true, one would also expect
-an interaction effect for AOM, upper respiratory tract
-infections, season, and number of siblings, or for a composite
-of all these, which was not found. The day-care interaction
effect therefore needs to be studied further in young children.

In the current analyses only European trials could be
included. A question could be raised whether the expected
benefits for VT for subgroups would be more evident in US
data. The main-effect results of individual trials performed in
the USA are very similar to those in the included trials—that
is, ventilation tubes had a beneficial effect on hearing in the
short term, but this effect disappeared in the long term.5,6
Paradise et al, for instance recently showed that prompt
insertion of ventilation tubes in children with otitis media
did not measurably improve developmental outcome in under 3
year olds.7 It is therefore not to be expected that inclusion of
data from the trials performed in the USA would have
changed the results of the meta-analysis.

Subgroups that might benefit more from treatment with
ventilation tubes include those with speech or language
-delays, behaviour and learning problems, Down’s syndrome,
or children with cleft palate. These could not be studied in
this IPD meta-analysis as these subgroups were excluded in
the individual trials. The experience of many clinicians that
these subgroups of children benefit more from treatment
with ventilation tubes has not yet been evidenced in RCTs. As
the question whether to treat these children with ventilation
tubes is very relevant for clinical practice, future trials
studying these specific subgroups are justified.

Prospective cohort studies have shown that OME during
early life may influence later language development nega-
tively,27,28 whereas others failed to find such an associa-
tion.29,30 These findings suggest that OME may not be an
innocent disease that should be left untreated. The insertion
of tubes can, on the other hand, lead to adverse effects such
as tympanosclerosis, atrophy, and retraction.31,32

In conclusion, the benefits of short term ventilation tubes
in children studied so far are small and related to the
continued presence and patency of the tubes. Therefore an
initial period of watchful waiting seems to be the appropriate
management strategy for most children with OME. None of

---

**Figure 2** (A) Stratified results regarding the interaction between day-
care and treatment (k = 4 trials) after 6 months follow up. p = 0.02 for
interaction. Number of cases in each subgroup: VT and day-care,
n = 123; WW + day-care, n = 118; VT without day-care, n = 67; WW
without day-care, n = 63. (B) Stratified results regarding the interaction
between day-care and treatment (k = 4 trials) after 12 months follow up.
p = 0.61 for interaction. Number of cases in each subgroup: VT and day-
care, n = 127; WW + day-care, n = 112; VT without day-care, n = 68;
WW without day-care, n = 61.

**Figure 3** (A) Stratified results regarding the interaction between
baseline hearing level and treatment in the trials that randomised ears
(k = 3) after 6 months follow up. p = 0.05 for interaction. Number of
cases in each subgroup: VT and ≤25 dB HL, n = 60; WW and ≤25 dB
HL, n = 57; VT and >25 dB HL, n = 20; WW and >25 dB HL, n = 23. (B)
Stratified results regarding the interaction between baseline hearing level
and treatment in the trials that randomised ears (k = 3) after 12 months
follow up. p = 0.28 for interaction. Number of cases in each subgroup:
VT and ≤25 dB HL, n = 58; WW and ≤25 dB HL, n = 55; VT and
>25 dB HL, n = 20; WW and >25 dB HL, n = 23.
the baseline variables showed a strong interaction effect with treatment justifying subgrouping. If ventilation tubes are used, they could be inserted in young children growing up in an environment with a high infection load (for example, children attending day-care) or in older children with a hearing level of 25 dB or greater in both ears persisting for at least 12 weeks. As no evidence is yet available for the subgroups of children with speech/language delays, behaviour and learning problems, and/or syndromes, the clinician will need to make his/her own decision regarding treatment for such children.

What is already known on this topic

- Trials performed so far have shown no or only marginal effects on hearing and language development.
- The insertion of ventilation tubes is, however, still the most common operation in children in western countries.

What this study adds

- Observation seems to be an adequate management strategy for most children with OME.
- Ventilation tubes might be used in young children that grow up in an environment with a high infection load (for example, children attending day-care), or in older children with a hearing level of 25 dB HL or greater in both ears, persisting for at least 12 weeks.

REFERENCES

Search strategy

We searched the Cochrane Central Register of Controlled Trials (CENTRAL) Cochrane Library Issue 1 2003, and we also searched PUBMED from 1966 to 2004 and EMBASE from 1973 to 2004. The date of the last search was June 2004. We searched for all RCTs and controlled clinical trials of surgical treatment. We used the terms “Otitis Media with Effusion”, “glue ear”, “serous otitis”, and “secretory otitis” as headings, and 8 separate headings for “grommets”, and “ventilation tubes”. To be selected for the IPD meta-analysis, trials had to be randomised to a high standard. Trials had to include children aged 0–12 years of age with tympanometrically and/or otoscopically confirmed persistent bilateral OME, and the comparison had to be between short term ventilation tubes (VT) and watchful waiting (WW).