RevMan: review – intervention; 6058210709132439 63 (version 1.7.262)

Status: UNPUBLISHED DRAFT

Ventilation tubes (grommets) for otitis media with effusion (OME) in children

Editors: Cochrane ENT Group

Contact Person: Samuel MacKeith (samuel.mackeith@ouh.nhs.uk)

ENT Department, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

Samuel MacKeith[1]Caroline A Mulvaney[2]Kevin Galbraith[2]Katie E Webster[3]

Rachel Connolly[4]Aye Paing[4]Tal Marom[5]Mat Daniel[6]Roderick P Venekamp[7]

Maroeska M Rovers[8]Anne GM Schilder[9][10]

[1] ENT Department, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

[2] Cochrane ENT, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK

[3] Cochrane ENT, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK

[4] Guideline Development Team A, NICE, London, UK

[5] Department of Otolaryngology-Head and Neck Surgery, Samson Assuta Ashdod University Hospital, Ben Gurion University, Ashdod, Israel

[6] Nottingham Children's Hospital, Nottingham, UK

[7] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, Netherlands

[8] Department of Operating Rooms, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands

[9] evidENT, Ear Institute, University College London, London, UK

[10] NIHR UCLH Biomedical Research Centre, University College London, London, UK

Citation

MacKeith S, Mulvaney CA, Galbraith K, Webster KE, Connolly R, Paing A, Marom T, Daniel M, Venekamp RP, Rovers MM, Schilder AGM. Ventilation tubes (grommets) for otitis media with effusion (OME) in children. Cochrane Database of Systematic Reviews TBD, Issue TBD. Art. No.: CD015215. DOI: 10.1002/14651858.CD015215.pub2.

Dates

Revision published: Issue TBD, TBD (TBD)

Version published (citation changed): Issue TBD, TBD (TBD)

Review first published: Issue TBD, TBD

Protocol first published: Issue 3, 2022

Abstract

Background

Otitis media with effusion (OME) is an accumulation of fluid in the middle ear cavity, common amongst young children. The fluid may cause hearing loss. When persistent, it may lead to developmental delay, social difficulty and poor quality of life. Management of OME includes watchful waiting, autoinflation, medical and surgical treatment. Insertion of ventilation tubes has often been used as the preferred treatment for this condition.

Objectives

To evaluate the benefits and harms of ventilation tubes for OME in children compared to no treatment, watchful waiting, myringotomy alone, hearing aids and other non-surgical treatment.

Search methods

The Cochrane ENT Information Specialist searched the Cochrane ENT Register; Central Register of Controlled Trials (CENTRAL); Ovid MEDLINE; Ovid Embase; Web of Science; ClinicalTrials.gov; ICTRP and additional sources for published and unpublished trials. The search date was 20 January 2023.

Selection criteria

We included randomised controlled trials (RCTs) and quasi-randomised trials in children (6 months to 12 years) with unilateral or bilateral OME for at least three months. We included studies that compared ventilation tube insertion with each of five comparators: no treatment, watchful waiting, myringotomy, hearing aids and other non-surgical treatments.

Data collection and analysis

We used standard Cochrane methods. Our primary outcomes were determined following a multi-stakeholder prioritisation exercise and were: 1) hearing; 2) OME-specific quality of life; 3) persistent tympanic membrane perforation (as a severe adverse effect of the surgery). Secondary outcomes were: 1) persistence of OME; 2) other adverse effects (including tympanosclerosis, ventilation tube blockage and pain); 3) receptive language skills; 4) speech development; 5) cognitive development; 6) psychosocial skills; 7) listening skills; 8) generic health-related quality of life; 9) parental stress; 10) vestibular function; 11) episodes of acute otitis media. We used GRADE to assess the certainty of evidence.

Although we included all measures of hearing assessment, the proportion of children who returned to normal hearing was our preferred method to assess hearing, due to challenges in interpreting the results of mean hearing thresholds.

Main results

We included 19 RCTs (2888 children). We assessed the data according to five main comparisons, described below. We considered most of the evidence to be very uncertain, due to wide confidence intervals for the effect estimates, relatively small numbers of participants, and a risk of performance and detection bias. Here we report our primary outcomes and main secondary outcome, at the longest reported follow-up. We did not identify data on disease-specific quality of life, however many of the studies were conducted before the development of otitis media-specific tools to assess quality of life.

Ventilation tubes compared to no treatment (four studies)

The odds ratio (OR) for a return to normal hearing after 12 months was 1.13 with ventilation tubes (95% confidence interval (CI) 0.46 to 2.74; 54% versus 51%; 1 study; 72 participants; very low-certainty evidence).

Ventilation tubes may result in a large reduction in persistence of OME at six months (risk ratio (RR) 0.30, 95% CI 0.14 to 0.65; 20% versus 68%; 1 study; 54 participants). At 12 months the OR was 0.66 (95% CI 0.24 to 1.85; 49% versus 58%; 1 study; 144 participants; very low-certainty evidence).

The evidence is very uncertain about the chance of tympanic membrane perforation with ventilation tubes (OR 0.85, 95% CI 0.38 to 1.91; 8.3% versus 9.7%; 1 RCT; 144 participants).

Early ventilation tubes compared to watchful waiting (ventilation tubes inserted later, if required) (six studies)

There was little difference in the proportion of children whose hearing returned to normal after 8 to 10 years (RR for ventilation tubes 0.98, 95% CI 0.94 to 1.03; 93% versus 95%; 1 study; 391 participants; very low-certainty evidence).

Ventilation tubes may also result in little difference in the risk of persistent OME after 18 months to 6 years (RR 1.21, 95% CI 0.84 to 1.74; 15% versus 12%; 3 studies; 584 participants; very low-certainty evidence).

Ventilation tubes compared to hearing aids

No studies considered this comparison.

Ventilation tubes compared to non-surgical treatment (one study)

One study assessed ventilation tubes compared to a six-month course of antibiotics (sulphisoxazole).

No data were available on return to normal hearing. The only evidence available considered final hearing thresholds. At four months the mean difference was -5.98 dB HL lower (better) for those receiving ventilation tubes (95% CI -9.21 to -2.75; 1 study; 125 participants; very low-certainty evidence).

Ventilation tubes compared to myringotomy (nine studies)

Ventilation tubes may slightly increase the likelihood of returning to normal hearing at 6 to 12 months, but the confidence intervals were very wide (RR 1.22, 95% CI 0.59 to 2.53; 74% versus 64%; 2 studies; 132 participants; very low-certainty evidence).

The evidence was also very uncertain about the persistence of OME after short- and medium-term follow-up, although the effect estimates tended to show a benefit from ventilation tubes. At long-term follow-up there may be little or no difference in the persistence of OME between those who received ventilation tubes and myringotomy (RR 0.97, 95% CI 0.90 to 1.05; 83% versus 85%; 1 study; 491 participants; low-certainty evidence).

Adverse effects across all comparisons

There is a risk of tympanic membrane perforation with ventilation tubes. We were unable to pool the data across different studies, but the absolute risk of perforation appears to be between 0% and 12%.

Authors' conclusions

When assessed with the GRADE approach, the evidence from RCTs for the use of ventilation tubes in OME is very uncertain. The evidence from the studies included does not allow us to say when (or how much) ventilation tubes improve hearing in any specific child. However, interpretation of the evidence is difficult: many children in the control groups recover spontaneously or receive ventilation tubes during the follow-up period, ventilation tubes may become blocked or fall out over time, and OME may recur.

For this reason, we do not believe that RCTs are necessarily the best way to determine whether a specific intervention is likely to be more effective than not in any specific child. Instead, we should first try to better understand the different OME phenotypes to target interventions to children who will benefit most, and avoid over-treating those who are likely to have spontaneous resolution of OME.

Plain language summary

Ventilation tubes (grommets) for glue ear in children

Key messages

From the studies included in this review, we are uncertain to what extent ventilation tubes improve hearing. Glue ear is a fluctuating condition, with high rates of spontaneous resolution and recurrence which makes it difficult to study in a clinical trial.

Ventilation tubes may slightly reduce the number of children who have glue ear after three to six months of follow-up. It is not clear whether they also have an effect over longer periods of time.

Insertion of ventilation tubes can lead to a persistent hole in the eardrum (tympanic membrane perforation), ranging from 0% to 12% in the studies that we assessed.

What is OME?

Glue ear (or 'otitis media with effusion', OME) is a relatively common condition affecting young children. Fluid collects in the middle ear, which may cause hearing impairment. As a result of their poor hearing, children may be behind in their speech and may have difficulties at school.

How is OME treated?

Most of the time OME does not need any treatment and the symptoms will get better with time. In children with persistent OME, different treatments have been used, including medications or surgery (insertion of grommets, with or without adenoidectomy). Ventilation tubes (grommets) are tiny plastic or silicon tubes, which are inserted in the eardrum under general anaesthesia. The tube allows fluid to drain out of the middle ear and allows air to enter.

What did we want to find out?

We wanted to identify whether insertion of ventilation tubes was better than no treatment, or other types of treatment (such as medicines or hearing aids), for children with OME.

We also wanted to see if there were any unwanted effects associated with having ventilation tubes inserted.

What did we do?

We searched for studies that compared ventilation tubes with either no treatment, or a different treatment, in children with OME. We compared and summarised the study results, and rated our confidence in the evidence, based on factors such as study methods and sizes.

What did we find?

We included 19 studies with a total of 2888 participants. We considered all the evidence we found to be uncertain, because of the relatively small number of children included and some issues with the conduct of the studies.

The evidence from the studies done so far does not allow us to say when, and by how much, ventilation tubes will improve hearing in any specific child.

Ventilation tubes may reduce the number of children with persistent OME after three to six months of follow-up. This benefit was not seen after longer follow-up. However, many children in the 'control group' (who planned to receive no treatment) either recovered spontaneously, or received ventilation tubes during the follow-up period. This makes it hard to assess the evidence after longer follow-up.

We did not find any evidence about quality of life, so we do not know if ventilation tubes have any impact on this.

We were not able to combine the results of different studies to calculate how often an eardrum perforation may occur. However, the studies reported this side effect in between 0% and 12% of people who received ventilation tubes.

What are the limitations of the evidence?

We did not have enough information to identify whether certain groups of children would benefit from ventilation tubes (for example, children with severe hearing loss, or those in a certain age group). Further work needs to be done to identify which children with OME would benefit from treatment, and which children are likely to recover spontaneously.

How up-to-date is this evidence?

The evidence is up-to-date to January 2023.

Summary of findings

Summary of findings 1

Ventilation tubes compared to no treatment for otitis media with effusion (OME) in children

Ventilation tubes compared to no treatment for otitis media with effusion (OME) in children						
Patient or population: children with otitis media with effusion (OME) **Setting:** outpatient **Intervention:** ventilation tubes **Comparison:** no treatment						
Outcomes	**Relative effect (95% CI)**	**Anticipated absolute effects* (95% CI)**			**Certainty of the evidence (GRADE)**	**What happens**
		Without ventilation tubes	**With ventilation tubes**	**Difference**		
Return to normal hearing Randomised by ear: normal defined as < 15 dB Assumed CC = 0.5 12 months (medium-term) № of participants: 72 (1 RCT)	OR 1.13 (0.46 to 2.74)	51.4%	54.4% (32.7 to 74.3)	3.0% more (18.7 fewer to 22.9 more)	⊕⊖⊖⊖ Very low[1]	The evidence is very uncertain about the effect of ventilation tubes on return to normal hearing at 12 months when compared with no treatment.
Persistence of OME, Randomised by child	RR 0.30 (0.14 to 0.65)	68.0%	20.4% (9.5 to 44.2)	47.6% fewer (58.5 fewer to 23.8 fewer)	⊕⊕⊖⊖ Low[2]	Ventilation tubes may result in a large reduction in the risk of persistence at 6 months when

Adjusted for non-independence of within-individual measurements, assumed ICC = 0.5 6 months (medium-term) № of participants: 54 (1 RCT)						compared with no treatment.
Persistence of OME Randomised by ear, assumed CC = 0.5 12 months (medium-term) № of participants: 144 (1 RCT)	OR 0.66 (0.24 to 1.85)	58.3%	48.0% (25.1 to 72.1)	10.3% fewer (33.2 fewer to 13.8 more)	⊕⊖⊖⊖ Very low[3]	The evidence is very uncertain about the effect of ventilation tubes on the likelihood of persistence at 12 months when compared with no treatment.
Adverse event: perforation/retraction Randomised by ear, assumed CC = 0.5 12 months (medium-term) № of participants: 144 (1 RCT)	OR 0.85 (0.38 to 1.91)	9.7%	8.4% (3.9 to 17.1)	1.3% fewer (5.8 fewer to 7.3 more)	⊕⊖⊖⊖ Very low[4]	The evidence is very uncertain about the effect of ventilation tubes on the likelihood of ear-drum perforation or retraction at 12 months when compared with no treatment.

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CC: correlation coefficient; **CI:** confidence interval; **ICC**: intracluster correlation coefficient; **OR:** odds ratio; **RR:** risk ratio

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]Downgraded by one level for a risk of performance bias. Downgraded by one level for inconsistency, as the I^2 was substantial (73%). Downgraded by one level for indirectness, as the definition of 'normal hearing' was particularly strict (< 15 dB). Downgraded by two levels for imprecision as the optimal information size (OIS) was not reached (< 300 events) and the confidence intervals crossed two decision thresholds (OR 0.80 and 1.25).

[2]Downgraded by one level for serious risk of performance and detection bias. Downgraded by one level for serious imprecision as the OIS was not reached (< 300 events).

[3]Downgraded by one level for serious risk of performance bias. Downgraded by two levels for imprecision as the optimal information size (OIS) was not reached (< 300 events) and the confidence intervals crossed two decision thresholds (OR 0.80 and 1.25).

[4]Downgraded by one level for a risk of performance bias. Downgraded by two levels for imprecision as the optimal information size (OIS) was not reached (< 300 events) and the confidence intervals crossed two decision thresholds (OR 0.80 and 1.25).

Summary of findings 2

Early ventilation tubes compared to watchful waiting (treatment later if required) for otitis media with effusion (OME) in children

Early ventilation tubes compared to watchful waiting (treatment later if required) for otitis media with effusion (OME) in children

Patient or population: children with otitis media with effusion (OME)
Setting: outpatient
Intervention: early ventilation tubes
Comparison: watchful waiting (treatment later if required)

Outcomes	Relative effect (95% CI)	Anticipated absolute effects* (95% CI)			Certainty of the evidence (GRADE)	What happens
		With watchful waiting	With early ventilation tubes	Difference		
Hearing returned to normal Randomised by child (age 9 to 11 - long-term) № of participants: 391 (1 RCT)	RR 0.98 (0.94 to 1.03)	94.9%	93.0% (89.2 to 97.7)	1.9% fewer (5.7 fewer to 2.8 more)	⊕⊖⊖⊖ Very low[1]	The evidence is very uncertain about the effect of early ventilation tubes on the return to normal hearing in the long term, when compared to watchful waiting (ventilation tubes later if required).
Presence/persistence of OME Randomised by child (1.5 to 9.75 years follow-up - long-term) № of participants: 584 (3 RCTs)	RR 1.21 (0.84 to 1.74)	12.2%	14.8% (10.3 to 21.3)	2.6% more (2 fewer to 9.1 more)	⊕⊖⊖⊖ Very low[2]	The evidence is very uncertain about the effect of early ventilation tubes on persistence of OME in the long term, when compared to watchful waiting (ventilation tubes later if required).
Adverse event: persistent perforation Follow-up: range 2 years to 3.75 years № of ears analysed: 1010 (2 RCTs)	One study (follow-up 3.75 years) yielded a RR for early ventilation tubes versus watchful waiting of 3.65 (95% CI 0.41, 32.38). One study (follow-up 2 years) reported that lasting perforations are rare and at worst 0.8% (5⁄635 ears that had ventilation tubes inserted).				⊕⊖⊖⊖ Very low[3]	The evidence is very uncertain about the effect of early ventilation tubes on the risk of persistent perforation when compared to watchful waiting (ventilation tubes later if required).

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; **RCT:** randomised controlled trial; **RR:** risk ratio

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]Downgraded by one level for serious risk of bias (performance bias), one level for serious indirectness (some children did not have a consecutive period of three months with OME before enrolment) and one level for serious imprecision (the optimal information size of 300 events was not reached).

[2]Downgraded by two levels for very serious risk of bias (due to performance bias and attrition bias), one level for serious indirectness (some children did not have a consecutive period of three months with OME before enrolment) and one level for serious imprecision (as the confidence interval crossed one decision threshold (RR 1.25)).

[3]Downgraded by one level for serious risk of bias (performance bias), one level for serious indirectness (some children did not have a consecutive period of three months with OME before enrolment) and one level for serious imprecision as a narrative synthesis was conducted, and no estimate of effect can be provided.

Summary of findings 3

Ventilation tubes compared to non-surgical treatment for otitis media with effusion (OME) in children

Ventilation tubes compared to non-surgical treatment for otitis media with effusion (OME) in children

Patient or population: children with otitis media with effusion (OME)
Setting: outpatient
Intervention: ventilation tubes
Comparison: non-surgical treatment

Outcomes	Relative effect (95% CI)	Anticipated absolute effects* (95% CI)			Certainty of the evidence (GRADE)	What happens
		With non-surgical treatment	With ventilation tubes	Difference		
Mean final hearing threshold (4 months - medium-term) № of participants: 125 (1 RCT)	—	The mean threshold without ventilation tubes was 17.8 dB	11.8 dB	MD 5.98 lower (9.21 lower to 2.75 lower)	⊕⊖⊖⊖ Very low[1]	The evidence is very uncertain about the effect of ventilation tubes on the hearing threshold at 4 months, when compared to non-surgical (antibiotic) treatment.
Adverse event: persistent perforation (18 months - long-term) № of participants: 60 (1 RCT)	One study reported that none of 60 children who received ventilation tubes had a persistent perforation. Length of follow-up was not reported directly, but assumed to be at the final examination at 18 months.				⊕⊕⊖⊖ Low[2]	Ventilation tubes may result in a low risk of persistent perforation at 18 months, when compared to non-surgical (antibiotic) treatment.

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).

CI: confidence interval; **MD:** mean difference; **RCT:** randomised controlled trial

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]Downgraded two levels for risk of bias, due to very serious risk of performance and detection bias. Downgraded one level for indirectness, as some children received a different (inferior) ventilation tube. Downgraded by one level for serious imprecision, as the optimal information size was not reached (400 participants).

[2]Not downgraded for risk of bias, as this outcome was felt to be sufficiently objective that it would not be impacted by performance or detection bias. Downgraded one level for indirectness, as some children received a different (inferior) ventilation tube. Downgraded by one level for serious imprecision, as this was a narrative synthesis only.

Summary of findings 4

Ventilation tubes compared to myringotomy for otitis media with effusion (OME) in children

Ventilation tubes compared to myringotomy for otitis media with effusion (OME) in children

Patient or population: children with otitis media with effusion (OME)
Setting: outpatient
Intervention: ventilation tubes
Comparison: myringotomy

Outcomes	Relative effect (95% CI)	Anticipated absolute effects* (95% CI)			Certainty of the	What happens
		With myringotomy	With ventilation	Difference		

			tubes		evidence (GRADE)	
Hearing returned to normal: ventilation tubes versus laser myringotomy (6 to 12 months - medium-term) Adjusted for non-independence of within-individual measurements. Assumed ICC of 0.5 № of participants: 132 (2 RCTs)	RR 1.22 (0.59 to 2.53)	63.6%	77.6% (37.5 to 100)	14.0% more (26.1 fewer to 97.4 more)	⊕⊖⊖⊖ Very low[1]	The evidence is very uncertain about the effect of ventilation tubes on the likelihood of a return to normal hearing at 6 to 12 months, when compared to laser myringotomy.
Persistence of OME: ventilation tubes versus thermal myringotomy, randomised by ear (3 months - short-term) № of participants: 72 (1 RCT)	Peto OR 0.11 (0.02 to 0.53)	19.4%	2.6% (0.5 to 11.3)	16.9% fewer (19 fewer to 8.1 fewer)	⊕⊖⊖⊖ Very low[2]	The evidence is very uncertain about the effect of ventilation tubes on persistence of OME at 3 months, when compared with thermal myringotomy.
Persistence of OME: ventilation tubes versus laser myringotomy (6 months - medium-term) Adjusted for non-independence of within-participant measurements: assumed ICC of 0.5 № of participants: 82 (1 RCT)	RR 0.32 (0.16 to 0.64)	49.0%	15.7% (7.8 to 31.4)	33.3% fewer (41.2 fewer to 17.6 fewer)	⊕⊖⊖⊖ Very low[3]	The evidence is very uncertain about the effect of ventilation tubes on persistence of OME at 6 months, when compared with laser myringotomy.
Persistence of OME: ventilation tubes versus laser myringotomy, randomised by ear (6 months - medium-term) Assumed CC of 0.5 № of participants: 272 (1 RCT)	OR 0.27 (0.19 to 0.38)	61%	29.7% (22.9 to 37.3)	31.3% fewer (38.1 fewer to 23.7 fewer)	⊕⊖⊖⊖ Very low[4]	The evidence is very uncertain about the effect of ventilation tubes on persistence of OME at 6 months, when compared with laser myringotomy.
Persistence of OME: ventilation tubes versus cold-steel myringotomy (2 years - long-term) № of participants: 491 (1 RCT)	RR 0.97 (0.90 to 1.05)	85.2%	82.7% (76.7 to 89.5)	2.6% fewer (8.5 fewer to 4.3 more)	⊕⊕⊖⊖ Low[5]	The evidence suggests that ventilation tubes results in little to no difference in the risk of persistent OME at 2 years, when compared with

						cold-steel myringotomy.
Persistent perforation Range of follow-up: 3 months to 2 years № of participants: 581 (6 RCTs)	The number of persistent perforations following insertion of ventilation tubes ranged from one ear to four ears, and from one to three children (D'Eredita 2006; Gates 1989; Tao 2020; To 1984). One study reported all thermal perforations were closed by 42 days (Ruckley 1988), and one study yielded a RR for persistent perforation (ventilation tubes versus laser myringotomy) of 1.0 (95% CI 0.06, 15.56) at 6 months (Yousaf 2016).				⊕⊕⊕⊖ Moderate[6]	Ventilation tubes likely increases the risk of persistent perforation. When compared with laser myringotomy, there is likely to be little to no difference in risk at 6 months.

*__The risk in the intervention group__ (and its 95% confidence interval) is based on the assumed risk in the comparison group and the __relative effect__ of the intervention (and its 95% CI).

CC: correlation coefficient; **CI:** confidence interval; **ICC**: intraclass correlation coefficient; **OR:** odds ratio; **RCT:** randomised controlled trial; **RR:** risk ratio

GRADE Working Group grades of evidence
High certainty: we are very confident that the true effect lies close to that of the estimate of the effect.
Moderate certainty: we are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.
Low certainty: our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.
Very low certainty: we have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[1]Downgraded two levels for risk of bias (performance and reporting bias). Downgraded one level for serious inconsistency, as the I^2 was 95%, with minimal overlap of confidence intervals. Downgraded two levels for very serious imprecision as the optimal information size (OIS) was not reached (< 300 events) and two decision thresholds were crossed by the CI (RR 0.80 and 1.25).

[2]Downgraded two levels for risk of bias (detection and reporting bias). Downgraded one level for serious imprecision as the optimal information size (OIS) was not reached (< 300 events).

[3]Downgraded two levels for risk of bias (performance and reporting bias). Downgraded one level for serious imprecision as the optimal information size (OIS) was not reached (< 300 events).

[4]Downgraded two levels for very serious risk of bias (performance, detection, reporting and attrition bias). Downgraded one level for serious imprecision as the optimal information size (OIS) was not reached (< 300 events).

[5]Downgraded two levels for very serious risk of bias (performance, detection and attrition bias).

[6]Not downgraded for risk of bias, as this outcome was felt to be sufficiently objective that it would not be impacted by performance or detection bias. Downgraded by one level for serious imprecision, as this was a narrative synthesis only.

Background

Description of the condition

Otitis media with effusion (OME) is a common condition in early childhood. The condition, also known as 'glue ear' and serous otitis media, is defined as "the presence of fluid in the middle ear without signs or symptoms of acute infection" (Rosenfeld 2016).

A key clinical feature of OME is hearing loss, due to decreased mobility of the tympanic membrane and consequent loss of sound conduction (Rosenfeld 2016). When hearing loss persists, this may affect speech and language development, and lead to behavioural problems in some children (NICE 2008). Other symptoms that may be attributable to OME include balance (vestibular) problems and ear discomfort (Rosenfeld 2016). When symptoms persist, they may lead to poor school performance and affect a child's daily activities, social interactions and emotions, possibly leading to a poorer quality of life for the child (Rosenfeld 2000).

It is thought that up to 80% of children have had OME by the age of four years but a decline in prevalence is observed for children beyond six years of age (Williamson 2011).

Most episodes of OME in children resolve spontaneously within three months, however approximately 35% of children will have more than one episode of OME and, furthermore, 5% to 10% of episodes will last for more than a year (Rosenfeld 2016). Children with OME following an episode of untreated acute otitis media (AOM) have a 59% rate of resolution by one month, rising to 74% by three months, while children with newly diagnosed OME of unknown duration demonstrate a resolution rate of 28% by three months and up to 42% by six months (Rosenfeld 2003). The condition is more prevalent in children with Down syndrome or cleft palate (Flynn 2009; Maris 2014). Atopy has been considered a potential risk factor for OME in children (Kreiner-Møller 2012; Marseglia 2008; Zernotti 2017).

Diagnosis of OME is typically by clinical examination including (pneumatic) otoscopy and/or tympanometry in primary care. Following diagnosis, there will often be a period of active observation, for at least three months. During the observation period the care provider may offer a non-surgical intervention such as hearing aids or autoinflation. NICE and AAO-HNS do not currently recommend the use of antibiotics, antihistamines, decongestants or corticosteroids for OME as there is insufficient evidence to suggest they are effective treatments (NICE 2008; Rosenfeld 2016). If OME has not resolved within the three-month observation period, the child may be referred for further management/active intervention. This may include hearing aid provision or review by an ENT surgeon for consideration for myringotomy, ventilation tubes insertion and/or adenoidectomy. The choice of active intervention varies considerably. Earlier active intervention may be considered for children at increased risk of developmental difficulties (see Rosenfeld 2016 for a list of 'at-risk' factors).

This Cochrane Review will focus on insertion of ventilation tubes as treatment for OME in children. This review forms part of a suite of five reviews of OME treatment that will address those interventions identified in a prioritisation exercise as being most important and in need of up-to-date Cochrane Reviews: namely, adenoidectomy, autoinflation, topical and oral steroids, and antibiotics (Cochrane ENT 2020).

Description of the intervention

NICE describes myringotomy and insertion of ventilation tubes (with or without adenoidectomy) as the most common surgical option for OME (NICE CKS 2021). Ventilation tubes (grommets) are tiny plastic tubes inserted in the tympanic membrane (under general anaesthetic in children). The procedure, undertaken by an ENT surgeon, involves making a small incision in the tympanic membrane (myringotomy), aspirating middle ear fluid as necessary and inserting the tube. The ventilation tube promotes middle ear ventilation and provides a passage for drainage of middle ear fluid. Generally, ventilation tubes eventually extrude into the external ear canal and the tympanic membrane closes (Venekamp 2018). In certain cases, early extrusion of the ventilation tubes occurs, and they may need replacing. While aspiration is common practice, there is little evidence to suggest that it is of benefit prior to ventilation tube insertion (Laina 2006).

Myringotomy can be performed alone without insertion of ventilation tubes, however when undertaken using 'cold steel' incision with a blade it results in rapid healing without maintenance of benefit. When undertaken using a laser to create a circular perforation in the tympanic membrane, healing and closure of the myringotomy perforation may take longer with more persisting benefits akin to a ventilation tube.

The role of adenoidectomy in addition to ventilation tubes has been assessed in a separate Cochrane Review (van den Aardweg 2010); this evidence will be updated as part of the new suite of five Cochrane Reviews of OME treatments and thus will not be assessed in this review.

How the intervention might work

For children with OME who suffer from hearing loss, the insertion of ventilation tubes helps the middle ear fluid to drain, aerates the middle ear and balances the pressures on each side of the tympanic membrane (Vanneste 2019), allowing for normal mobility and

conduction of sound and thus improving the child's ability to hear. The improvement in hearing is immediate in the majority of cases but occasionally complete resolution takes days to weeks. Ventilation tubes usually remain working within the tympanic membrane for 12 months on average (Rosenfeld 2016), and usually spontaneously extrude with healing of the tympanic membrane. Following this the child may remain free from OME, however in a proportion of children OME can return and persist, requiring repeat insertion. Factors that can limit the effectiveness of ventilation tubes include blockage of the tube (with blood), difficulty or inability to place the tubes due to narrow ear canals (Down syndrome and cleft palate) and early extrusion.

A common problem with ventilation tubes is ear discharge (otorrhoea) (Schilder 2016), and in around 2% of cases when the ventilation tube is extruded the tympanic membrane does not heal and a perforation persists. There is some evidence that insertion of ventilation tube may also result in long-term damage to the tympanic membrane, such as tympanosclerosis or atrophy, and hearing loss (de Beer 2004; de Beer 2005).

Why it is important to do this review

A Cochrane Review assessing ventilation tubes for hearing loss associated with OME was published in 2010 (Browning 2010), updating an earlier review published in 2005. The 2010 review included 10 studies, three of which were randomised by ear (unilateral ventilation tube) and seven were randomised by child (bilateral ventilation tube or no ventilation tube). The authors concluded that the effect of ventilation tubes on hearing was small and diminished after six to nine months (by which time the hearing of children without ventilation tubes had improved due to natural resolution). The authors found few data on other outcomes, and identified a lack of trials conducted in children with established speech, language, learning or developmental problems. Since publication of the Cochrane Review in 2010 there have been two Health Technology Assessment (HTA) reports that include ventilation tubes (Berkman 2013; Steele 2017), and four other systematic reviews (Berkman 2013; Cheong 2012; Wallace 2014; Williamson 2011). Scoping searches for randomised controlled trials (RCTs) of ventilation tubes, which were last undertaken in January 2020, identified 12 abstracts of interest published since the last Cochrane Review. A prioritisation exercise undertaken in 2020 identified a review of ventilation tubes as a top priority (Cochrane ENT 2020). It is therefore timely to update the evidence.

Objectives

To assess the effects (benefits and harms) of ventilation tubes (grommets) for OME in children.

Methods

Criteria for considering studies for this review

Types of studies

We included randomised controlled trials (RCTs) and quasi-randomised trials (where studies were designed as RCTs, but the sequence generation for allocation of treatment used methods such as alternative allocation, birth dates and alphabetical order). We included studies that randomised participants by ear, by participant or by cluster. We did not identify any cluster-randomised or cross-over trials for inclusion in this review.

Types of participants

The population of interest is children aged 6 months to 12 years with unilateral or bilateral otitis media with effusion, alternatively termed chronic otitis media with effusion (COME), glue ear, chronic or persistent middle-ear effusion or serous otitis media. If a study included children aged younger than 6 months and/or older than 12 years, we only

included the study if the majority of children fit our inclusion criteria, or if the trialists presented outcome data by age group. We included all children regardless of any comorbidity such as Down syndrome or cleft palate.

Clinical diagnosis of OME was confirmed by oto(micro)scopy or tympanometry or both. We included studies where children had OME for at least three months. We included studies of children who had previously had ventilation tubes inserted.

In some studies, the population of interest was children with acute otitis media (AOM) or recurrent acute otitis media (RAOM). Either of these populations may also have intermittent or chronic OME. However, we regarded children who present with AOM or RAOM as different populations to those who present with chronic OME (the focus of this review), and did not assume that interventions designed to treat one population will have the same efficacy in the others. We therefore excluded studies in which the population of interest was children with AOM or RAOM.

Types of interventions

Intervention

Insertion of ventilation tube performed either unilaterally or bilaterally. We did not assess different types of ventilation tubes or surgical approaches to insertion.

Comparator

In our protocol we presented six comparisons of interest. However, after examining the comparisons of interest it was agreed that the comparisons of 'no treatment' and 'watchful waiting' are not the same and should not be treated as one comparison. The comparison of watchful waiting requires an active process of monitoring the child's condition and treating them with the intervention, such as bilateral VT, if deemed necessary at a later date.

As some studies included children with both bilateral and unilateral OME it was also decided to merge those comparisons where trials might include these participants. Hence we are interested in the following five comparisons:

- ventilation tubes (bilateral or unilateral) versus no treatment
- early ventilation tubes versus watchful waiting (treatment later if required);
- ventilation tubes versus hearing aids;
- ventilation tubes versus non-surgical treatment;
- ventilation tubes versus myringotomy alone.

If study participants received other treatments (for example, adenoidectomy, intranasal steroids, oral steroids, antibiotics, mucolytics or decongestants) we included these studies if both arms received identical treatment.

Types of outcome measures

We analysed the following outcomes in the review, but we will not use them as a basis for including or excluding studies. We assessed all outcomes in the short term (≤ 3 months), medium term (> 3 months to ≤ 1 year) and long term (> 1 year). We assessed postoperative adverse events in the very short term (< 6 weeks).

Primary outcomes

- Hearing, measured as:
    - the proportion of children whose hearing has returned to normal (defined by the trialists);
    - mean final hearing threshold (determined for the child or ear, depending on the unit of analysis);

- change in hearing threshold from baseline (determined for the child or ear, depending on the unit of analysis).

We anticipated that trial data for these outcomes may be derived from a variety of assessment methods and subject to a variety of definitions. To avoid loss of important evidence, we extracted all such data for analysis. However, we gave consideration to the appropriateness of pooling different types of data in meta-analysis. Our selection of primary outcomes was based principally upon clinical importance, but also permits applicability across a variety of age-appropriate assessment methods, and considers the types of outcome data that are most likely to be available. Accordingly, we regarded the proportion of participants whose hearing has returned to normal as the most important measure of hearing impact. We considered medium- and long-term outcome data as the most clinically important.

- Disease-specific quality of life measured using a validated instrument, for example:
  - OM8-30 (Haggard 2003);
  - Otitis Media-6 (Rosenfeld 1997).
- Adverse event - persistent perforation.

Secondary outcomes

- Presence/persistence of OME.
- Adverse events - measured by the number of participants affected.
  - Tympanic membrane changes, such as:
    - atrophy;
    - atelectasis or retraction;
    - myringosclerosis;
    - tympanosclerosis.
  - Tube-related, such as:
    - blockage;
    - extrusion;
    - granulation tissue formation;
    - otorrhoea/perforation;
    - displacement of the ventilation tube into the middle ear space.
  - Patient-related, such as:
    - vomiting;
    - diarrhoea;
    - dry throat;
    - nasal stinging;
    - cough;
    - long-term hearing loss;
    - postsurgical haemorrhage;
    - pain.
- Receptive language skills, measured using a validated scale, for example:
  - Peabody Picture Vocabulary Test - Revised (Dunn 2007);
  - relevant domains of the Reynell Developmental Language Scales (Reynell 1985);
  - relevant domains of the Preschool Language Scale (PLS) (Zimmermann 1992);

    - relevant domains of the Sequenced Inventory of Communication (SCID) (Hedrick 1984).
- Speech development, or expressive language skills, measured using a validated scale, for example:
    - Schlichting test (Schlichting 2010);
    - Lexi list (Schlichting 2007);
    - relevant domains of the Reynell Developmental Language Scales (Reynell 1985);
    - relevant domains of the PLS (Zimmermann 1992);
    - relevant domains of the SCID (Hedrick 1984).
- Cognitive development, measured using a validated scale, for example:
    - Griffiths Mental Development Scales (Griffiths 1996);
    - McCarthy General Cognitive Index (McCarthy 1972);
    - Bayley Scales of Infant and Toddler Development (Bayley 2006).
- Psychosocial outcomes, measured using a validated scale, for example:
    - the Social Skills Scale of the Social Skills Rating System (Gresham 1990);
    - Child Behavior Checklist (Achenbach 2011);
    - Strengths and Difficulties Questionnaire (Goodman 1997);
    - Pediatric Symptom Checklist (Jellinek 1988).
- Listening skills, for example listening to stories and instructions effectively. Given that there are few validated scales to assess listening skills in children with OME, we will include any methods used by trialists.
- Generic health-related quality of life assessed using a validated instrument, for example:
    - EQ-5D (Rabin 2001);
    - TNO AZL Children's QoL (TACQOL) (Verrips 1998);
    - TNO AZL Pre-school children QoL (TAPQOL) (Fekkes 2000);
    - TNO AZL Infant Quality of Life (TAIQOL) (TNO 1997);
    - Infant Toddler Quality of Life Questionnaire (ITQOL) (Landgraf 1994);
    - Child Health Questionnaire (CHQ) (Landgraf 1996).
- Parental stress, measured using a validated scale, for example:
    - Parenting Stress Index (Abidin 1995).
- Vestibular function:
    - balance;
    - co-ordination.
- Number of doctor-diagnosed AOM episodes within a specified time frame.

These outcomes were identified as the most important in two studies that aimed to develop a core outcome set for children with OME (Bruce 2015; Liu 2020). As this review forms part of a suite of reviews of interventions for OME, not all outcomes will be relevant for all reviews.

Search methods for identification of studies

The Cochrane ENT Information Specialist conducted systematic searches for randomised controlled trials and controlled clinical trials. There were no language, publication year or publication status restrictions. We contacted original authors for clarification and further

data if trial reports were unclear and arranged translations of papers where necessary. The date of the search was 20 January 2023.

Electronic searches

The Information Specialist searched:

- the Cochrane ENT Register (searched via the Cochrane Register of Studies to 20 January 2023);
- the Cochrane Central Register of Controlled Trials (CENTRAL) (searched via the Cochrane Register of Studies to 20 January 2023);
- Ovid MEDLINE(R) Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Ovid MEDLINE(R) Daily and Ovid MEDLINE(R) (1946 to 20 January 2023);
- Ovid EMBASE (1974 to 20 January 2023);
- Web of Science, Web of Science (1945 to 20 January 2023);
- ClinicalTrials.gov, www.clinicaltrials.gov:
    - searched via the Cochrane Register of Studies to 20 January 2023;
    - searched via www.clinicaltrials.gov to 20 January 2023;
- World Health Organization (WHO) International Clinical Trials Registry Platform (ICTRP), https://apps.who.int/trialsearch/:
    - searched via the Cochrane Register of Studies to 20 January 2023;
    - searched via https://apps.who.int/trialsearch/ 20 January 2023.

The Information Specialist modelled subject strategies for databases on the search strategy designed for CENTRAL. The search strategies were designed to identify all relevant studies for a suite of reviews on various interventions for otitis media with effusion. Where appropriate, they were combined with subject strategy adaptations of the highly sensitive search strategy designed by Cochrane for identifying randomised controlled trials and controlled clinical trials (as described in the Technical Supplement to Chapter 4 of the *Cochrane Handbook for Systematic Reviews of Interventions* version 6.1) (Lefebvre 2020). Search strategies for major databases including CENTRAL are provided in Appendix 1.

Searching other resources

We scanned the reference lists of identified publications for additional trials and contacted trial authors where necessary. The Information Specialist also ran non-systematic searches of Google Scholar to retrieve grey literature and other sources of potential trials.

We did not perform a separate search for adverse effects. We considered adverse effects described in included studies only.

Data collection and analysis

Selection of studies

The Cochrane ENT Information Specialist used Cochrane's Screen4Me workflow to help assess the search results. Screen4Me comprises three components:

1. Known assessments – a service that matches records in the search results to records that have already been screened in Cochrane Crowd and been labelled as 'a RCT' or as 'not a RCT'.
2. The machine learning classifier (RCT model) (Wallace 2017), available in the Cochrane Register of Studies (CRS-Web), which assigns a probability of being a true RCT (from 0 to 100) to each citation. For citations that are assigned a probability score below the cut-point at a recall of 99% we will assume these to be

non-RCTs. For those that score on or above the cut-point we will either manually dual screen these results or send them to Cochrane Crowd for screening.

3. Cochrane Crowd is Cochrane's citizen science platform where the Crowd help to identify and describe health evidence. For more information about Screen4Me and the evaluations that have been done, please go to the Screen4Me website on the Cochrane Information Specialist's portal and see Marshall 2018, McDonald 2017, Noel-Storr 2018 and Thomas 2017.

Two review authors (KG, CM) independently screened the remaining titles and abstracts to identify potentially relevant studies. At least two review authors (of KG, SM, CM and KW) then independently evaluated the full text of each potentially relevant study to determine whether it met the inclusion/exclusion criteria for this review. Any differences were resolved by discussion and consensus, with the involvement of a third author where necessary.

Screening eligible studies for trustworthiness

Two review authors appraised all studies meeting our inclusion criteria for trustworthiness using a screening tool developed by Cochrane Pregnancy and Childbirth. This tool includes specified criteria to identify studies that are considered sufficiently trustworthy to be included in the review (see Appendix 2 and Figure 1). For any studies assessed as being potentially 'high risk', we attempted to contact the study authors to obtain further information or address any concerns. We had planned to exclude these studies from the review if we were unable to contact the authors, or there was persisting uncertainty about the study.

When using the trustworthiness tool, there were 11 studies where we had no concerns: Bernard 1991; Gates 1989; Koopman 2004; Maw 1983; Maw 1999; Paradise 2007; Rach 1991; Rovers 2000; Ruckley 1988; TARGET 2000; To 1984.

All of the remaining studies had at least some concerns, although this was often due to a paucity of information, rather than a specific concern over trustworthiness:

- We were unable to identify prospective trial registration for six studies (Elkholy 2021; Popova 2010; Sujatha 2015; Tao 2020; Velepic 2011; Yousaf 2016).
- Four studies reported full follow-up, without explanation to indicate how this was achieved (Elkholy 2021; Sujatha 2015; Velepic 2011; Yousaf 2016).
- Three studies randomised equal numbers of participants to each group, without a description of blocked randomisation (D'Eredita 2006; Elkholy 2021; Sujatha 2015) and one did not provide information on the number randomised to each group (Dempster 1993).

We were unsure whether the number of studies with concerns reflected a genuine problem with the data from these studies, or whether the assessment tool was perhaps too sensitive. We note that this tool - and others used for the same purpose - has not yet been validated.

Consequently we decided to include all of the studies in the main analyses of this review, but we did investigate the effect of excluding studies with concerns over trustworthiness on the overall results (see Sensitivity analysis).

Data extraction and management

Two review authors (of RC, KG, CM, AP and KW) independently extracted outcome data from each study using a standardised data collection form. Where a study had more than one publication, we retrieved all publications to ensure complete extraction of data. Any discrepancies in the data extracted by the two authors were checked against the original reports, and differences were resolved through discussion and consensus, with recourse to a third author where necessary. If required, we contacted the study authors for clarification of any unclear or missing data. We included key characteristics of the studies, such as the study design, whether randomised by individual or by body part (see Unit of

analysis issues), setting, sample size, population and the methods for defining or collecting outcome data in the studies.

We extracted data on study findings according to treatment assignment, irrespective of whether study participants complied with treatment or received the treatment to which they were randomised.

In addition to extracting pre-specified information about study characteristics and aspects of methodology relevant to risk of bias, we extracted the following summary statistics for each trial and outcome:

- For continuous data: the mean values, standard deviation and number of patients for each treatment group at the different time points for outcome measurement. Where endpoint data were not available, we extracted the values for change-from-baseline data instead. If values for the individual treatment groups were not reported, where possible we extracted summary statistics (e.g. mean difference) from the studies.
- For binary data: we extracted information on the number of participants experiencing an event, and the number of participants assessed at that time point. If values for the individual treatment groups were not reported, where possible we extracted summary statistics (e.g. risk ratio) from the studies.
- For ordinal scale data: if the data appeared to be normally distributed, or if the analysis performed by the investigators indicated that parametric tests were appropriate, then we treated the outcome measure as continuous data. Alternatively, if data were available, we converted these to binary data for analysis.

We pre-specified time points of interest for the outcomes in this review. Where studies reported data at multiple time points, we took the longest available follow-up point within each of the specific time frames. For example, if a study reported an outcome at 4 months, 8 months and 12 months of follow-up then the 12-month data was included for the time point > 3 months to ≤ 1 year. For adverse events, some studies reported frequency data for events and it may not be possible to determine whether these events occurred in one participant on one occasion or more than one occasion. In such circumstances we will report the data narratively.

Assessment of risk of bias in included studies

Two authors (of RC, KG, CM, AP and KW) undertook assessment of the risk of bias of the included studies independently, with the following taken into consideration, as guided by the *Cochrane Handbook for Systematic Reviews of Interventions* (Handbook 2011):

- sequence generation;
- allocation concealment;
- blinding of participants and personnel;
- blinding of outcome assessment;
- incomplete outcome data;
- selective outcome reporting;
- other sources of bias.

We used the Cochrane risk of bias tool in RevMan 5.4 (RevMan 2020), which involves describing each of these domains as reported in the study and then assigning a judgement about the adequacy of each entry: 'low', 'high' or 'unclear' risk of bias.

Measures of treatment effect

We summarised dichotomous data, such as presence of OME, as risk ratios (RR) and 95% confidence intervals (CI) and we summarised continuous data as mean difference (MD) and 95% CI. For the outcomes presented in the summary of findings tables, we have provided both relative and absolute measures of effect. If individual patient data

(IPD) were available we planned to use these in our analyses, however, this was not possible.

Unit of analysis issues

Studies included in this review randomised either by participant, or by ear. We identified whether randomisation was conducted at the level of the participant or the ear, and - for those studies that randomised by participant - we assess whether the study included one or two ears from each participant. Given that there are likely to be some carry-over effects of disease and treatment from one ear to the other in a child, we analysed the outcomes separately for randomisation by ear or by child. For studies that randomised by ear, we only assessed the outcomes of hearing, adverse events, presence of OME and number of AOM episodes. The remaining outcomes are only relevant for studies randomised by child, where we can consider the more global effect of hearing difficulty.

Dealing with missing data

We attempted to contact study authors by email where data on an outcome of interest to the review were not reported but the methods described in the paper suggested that the outcome was assessed. We did the same if not all data required for meta-analysis were reported.

Assessment of heterogeneity

We assessed clinical heterogeneity by examining the included studies for potential differences in the types of participants recruited, interventions or controls used, and the outcomes measured. We assessed statistical heterogeneity by considering both the I^2 statistic (which calculates the percentage of variability that is due to heterogeneity rather than chance, with values over 50% suggesting substantial heterogeneity) and the P value from the Chi^2 test (Higgins 2021).

Assessment of reporting biases

We assessed reporting bias as within-study outcome reporting bias and between-study publication bias.

Outcome reporting bias (within-study reporting bias)

We assessed within-study reporting bias by comparing the outcomes reported in the published report against the study protocol or trial registry, when this could be obtained. If the protocol or trial registry entry was not available, we compared the outcomes reported to those listed in the methods section of the published report. If results were mentioned but not reported in a way that allows analysis (e.g. the report only mentions whether the results were statistically significant or not), we sought further information from the study authors. If no further information could be found, we noted this as being a 'high' risk of bias. If there was insufficient information to judge the risk of bias we noted this as an 'unclear' risk of bias (Handbook 2011).

Publication bias (between-study reporting bias)

We planned to produce a funnel plot to explore possible publication biases, if we were able to pool 10 or more studies in a single analysis. However, this was not possible, as too few studies were included in the meta-analyses.

Data synthesis

Where two or more studies reported the same outcome we performed a meta-analysis using Review Manager 5 (RevMan 2020). We report pooled effect measures for dichotomous outcomes as a risk ratio (RR) using the Mantel-Haenszel methods. For continuous outcomes measured using the same scales we report the mean difference (MD). We used a random-effects model.

Where it was not possible to pool the findings from studies in a meta-analysis, we have presented the results of each study and provide a narrative synthesis of findings.

Subgroup analysis and investigation of heterogeneity

We planned to analyse the following subgroups if sufficient data were available in study reports:

- children with mild hearing loss versus moderate or worse;
- children with allergy versus those without (using the trialists' own definition);
- children aged up to four years versus children aged four years and over;
- children with previous ventilation tubes versus those without ventilation tubes;
- children with cleft palate versus children without;
- children with Down syndrome versus children without;
- conventional cold steel versus other methods of myringotomy.

However, we did not find any data suitable for conducting these subgroup analyses. No studies provided subgroup data for children with different features (for example, for those with mild hearing loss, compared to those with moderate or worse hearing loss). Many of the studies did not provide sufficient background information (for example, on hearing level) for us to conduct subgroup analysis at the level of the individual study. Although we identified some studies that specifically recruited children aged up to four years or over four years, we had too few studies included in any meta-analysis to provide accurate estimates of subgroup effects.

Sensitivity analysis

We carried out sensitivity analyses to assess whether our findings were robust to decisions made regarding the analyses and inclusion of studies. We performed sensitivity analyses to assess the following:

- impact of model chosen: we compared the results using a random-effects versus a fixed-effect model;
- inclusion of studies at high risk of risk of bias: we compared the results including all studies versus excluding studies at overall high risk of bias, that is four or more of the seven domains of bias are rated as high risk (see Assessment of risk of bias in included studies). This applied to six studies (Elkholy 2021; Gates 1989; Koopman 2004; Popova 2010; Velepic 2011; Yousaf 2016);
- exclusion of studies with concerns over trustworthiness, as assessed by the Trustworthiness Tool (Figure 1). This applied to eight studies (D'Eredita 2006; Dempster 1993; Elkholy 2021; Popova 2010; Sujatha 2015; Tao 2020; Velepic 2011; Yousaf 2016).

The results of these analyses are presented in Table 1.

Summary of findings and assessment of the certainty of the evidence

Two independent authors (KG, CM) used the GRADE approach to rate the overall certainty of evidence using GRADEpro GDT (https://gradepro.org/). The certainty of evidence reflects the extent to which we are confident that an estimate of effect is correct, and we have applied this in the interpretation of results. There are four possible ratings: high, moderate, low and very low. A rating of high certainty of evidence implies that we are confident in our estimate of effect and that further research is very unlikely to change our confidence in the estimate of effect. A rating of very low certainty implies that any estimate of effect obtained is very uncertain.

The GRADE approach rates evidence from RCTs that do not have serious limitations as high certainty. However, several factors can lead to the downgrading of the evidence to moderate, low or very low. The degree of downgrading is determined by the seriousness of these factors:

- study limitations (risk of bias);

- inconsistency;
- indirectness of evidence;
- imprecision; and
- publication bias.

When assessing imprecision, we used a minimally important difference of a risk ratio (or odds ratio) of 0.8 or 1.25 for dichotomous outcomes. For most continuous data we considered a minimally important difference to be half of the standard deviation for the control/comparator group. The exception to this was hearing thresholds, where a difference of 10dB HL was used as the minimally important difference.

We include a summary of findings table, constructed according to the recommendations described in Chapter 14 of the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2021), for the following comparisons:

- ventilation tubes (bilateral or unilateral) versus no treatment;
- early ventilation tubes versus watchful waiting (treatment later if required);
- ventilation tubes versus hearing aids;
- ventilation tubes versus non-surgical treatment;
- ventilation tubes versus myringotomy alone.

We included the following four outcomes in the summary of findings table:

- hearing;
- disease-specific quality of life;
- presence/persistence of OME;
- adverse event - persistent perforation.

Results

Description of studies

Results of the search

The searches (January 2023 and September 2021) retrieved a total of 7441 records. This reduced to 4157 after the removal of duplicates. The Cochrane ENT Information Specialist sent all 4157 records to the Screen4Me workflow. The Screen4Me workflow identified 68 records as having previously been assessed: 50 had been rejected as not RCTs and 34 had been assessed as possible RCTs. The RCT classifier rejected an additional 1514 records as not RCTs (with 99% sensitivity). The Cochrane Crowd assessed the remaining 2443 references, rejecting 1313 as not RCTs and identifying 1130 as possible RCTs. Following this process, the Screen4Me workflow had rejected 2877 records and identified 1280 possible RCTs for title and abstract screening.

	Possible RCTs	Rejected
Known assessments	34	50
RCT classifier	2559	1514
Cochrane Crowd	1130	1313
Total	1280	2877

We identified 76 additional duplicates. We screened the titles and abstracts of the remaining 1204 records. We discarded 886 records and assessed 318 full-text records. We subsequently discarded an additional 192 records and identified an additional five duplicates.

We excluded 50 records (linked to 47 studies) with reasons recorded in the review (see Excluded studies).

We included 19 studies (64 records) where results were available (Bernard 1991; D'Eredita 2006; Dempster 1993; Elkholy 2021; Gates 1989; Koopman 2004; Maw 1983; Maw 1999; Paradise 2007; Popova 2010; Rach 1991; Rovers 2000; Ruckley 1988; Sujatha 2015; Tao 2020; TARGET 2000; To 1984; Velepic 2011; Yousaf 2016).

We identified three ongoing studies. See Characteristics of ongoing studies for further details.

We identified four studies that remain in awaiting assessment because we did not have enough information to determine eligibility (Diacova 2016; Marshak 1980; Maw 1986; Tawfik 2002).

A flow chart of study retrieval and selection is provided in Figure 2.

Included studies

A full description of each study is available in Characteristics of included studies, and a summary across all studies can be seen in Table 2

Study design

All the included studies were described as randomised controlled trials. Most were parallel group studies including two arms (Bernard 1991; D'Eredita 2006; Elkholy 2021; Maw 1999; Paradise 2007; Popova 2010; Rach 1991; Rovers 2000; Sujatha 2015; Tao 2020; Velepic 2011; Yousaf 2016). The TARGET 2000 study included a third arm, but these data were not relevant for this review (as they assessed adenoidectomy).

Three further studies were also 2-arm trials, but recruited children with bilateral OME - one ear of each child was assigned to the intervention, and the other ear was assigned to the comparator group (Koopman 2004; Ruckley 1988; To 1984).

Three studies with 4-arms were included. One compared ventilation tubes to myringotomy, and ventilation tubes plus adenoidectomy to adenoidectomy alone (Gates 1989). The two further studies randomised children with bilateral OME to adenoidectomy or no adenoidectomy, then assigned different interventions to each ear (Dempster 1993; Maw 1983). For the purposes of this review we have only made a comparison of those who received ventilation tubes to no ventilation tubes.

Location

Six studies were conducted in the UK (Dempster 1993; Maw 1983; Maw 1999; Ruckley 1988; TARGET 2000; To 1984), three in the USA (Bernard 1991; Gates 1989; Paradise 2007) and three in the Netherlands (Koopman 2004; Rach 1991; Rovers 2000). A single study was conducted in each of the following countries: Bulgaria (Popova 2010), China (Tao 2020), Croatia (Velepic 2011), Egypt (Elkholy 2021), India (Sujatha 2015), Italy (D'Eredita 2006) and Pakistan (Yousaf 2016).

Participants

Sample size

The size of the studies varied considerably, with the smallest study including only 30 participants (D'Eredita 2006). Nine studies recruited between 40 and 100 participants (Dempster 1993; Elkholy 2021; Maw 1983; Popova 2010; Rach 1991; Ruckley 1988; Sujatha 2015; To 1984; Velepic 2011; Yousaf 2016) and six studies included between 100 and 250 participants (Bernard 1991; Koopman 2004; Maw 1999; Rovers 2000; Tao 2020; TARGET 2000). Only two studies recruited more than 250 participants: Gates 1989 (578 subjects) and Paradise 2007 (429 subjects).

Age

Four studies recruited very young children:

- Paradise 2007 included children aged less than 3 years
- Maw 1999 included children aged between 9 months and 4 years
- Rach 1991 included children aged 2-4 years with bilateral OME

- Rovers 2000 included children who had failed a routine hearing screening test at the age of 9 months, and subsequently failed follow-up tests. The mean age of participants at recruitment was 19.5 months.

Most studies recruited slightly older children, typically aged between 3 and 12 years of age (Bernard 1991; D'Eredita 2006; Dempster 1993; Elkholy 2021; Gates 1989; Koopman 2004; Ruckley 1988; Sujatha 2015; Tao 2020; TARGET 2000; To 1984; Yousaf 2016). Three studies did not give age restrictions as part of their inclusion criteria, but the baseline characteristics of the participants indicated that the mean age was approximately 5-6 years (Maw 1983; Popova 2010; Velepic 2011).

Hearing loss

Many of the studies required participants to have confirmed hearing loss on entry to the trial. However, the requirements varied considerably.

- One study recruited children who failed a hearing test - with no response to sounds presented at 35dB (Rovers 2000)
- One study required a hearing level of more than 30dBHL (Yousaf 2016)
- Five studies included children with a hearing of at least 25dBHL (Bernard 1991; Dempster 1993; Maw 1983; Maw 1999; Tao 2020)
- Two studies recruited children with hearing loss of >20dBHL (Popova 2010; TARGET 2000)
- One study stated that the air-bone gap should be at least 25dB (Sujatha 2015)
- One study required parents to have noticed impaired hearing, but did not use a specific threshold for recruitment (Koopman 2004).

Eight studies did not explicitly state the level of hearing impairment which was necessary for enrolment in the study (D'Eredita 2006; Elkholy 2021; Gates 1989; Paradise 2007; Rach 1991; Ruckley 1988; To 1984; Velepic 2011).

Previous treatment

Most studies specifically excluded individuals who had previous received ventilation tubes and/or adenoidectomy (Bernard 1991; D'Eredita 2006; Dempster 1993; Elkholy 2021; Gates 1989; Tao 2020; TARGET 2000; To 1984; Velepic 2011). Some children enroled in the study by Koopman 2004 had previously undergone adenoidectomy, ventilation tube insertion or tonsillectomy.

A few studies specifically recruited children who had failed some form of medical therapy - typically antibiotics, with or without decongestants (Bernard 1991; Elkholy 2021; Gates 1989; Sujatha 2015; Paradise 2007), whilst two studies recruited children early in their presentation with OME, although it was not clear whether they may have received some form of medical therapy at presentation (Ruckley 1988; TARGET 2000).

No information on previous treatment was provided by six studies (Maw 1983; Maw 1999; Popova 2010; Rach 1991; Rovers 2000; Yousaf 2016).

Other health issues

The majority of studies specifically excluded children with congenital risk factors for OME, including cleft palate and Down syndrome (Bernard 1991; D'Eredita 2006; Dempster 1993; Elkholy 2021; Gates 1989; Maw 1999; Popova 2010; Rach 1991; Rovers 2000; Sujatha 2015; Tao 2020; TARGET 2000; Velepic 2011).

Interventions and comparisons

Comparison 1: Ventilation tubes versus no treatment

We identified four studies for this comparison. Two studies compared outcomes within the same individual - comparing insertion of a ventilation tube in one ear, to no surgery on the other ear (Dempster 1993; Maw 1983). One study compared outcomes for bilateral ventilation tube insertion (in both ears of the same individual) to no treatment (in other

children) (Rach 1991). For the study by Elkholy 2021, randomisation was also at the level of the individual child, but we were uncertain whether children received bilateral or unilateral ventilation tubes.

Children in Dempster 1993 were also randomised to receive adenoidectomy or no adenoidectomy. For this review, we have presented data separately (for those who did or did not receive adenoidectomy), but have also presented a pooled estimate of the overall effect of ventilation tube insertion. All children recruited to Elkholy 2021 also received adenoidectomy.

In the study by Rach 1991, randomisation was by child, but the individual ear was the unit of analysis for persistence of OME - results have therefore been adjusted to account for the correlation between ears of the same individual.

Comparison 2: Ventilation tubes versus watchful waiting

This comparison included six studies where some children were randomised to receive ventilation tubes immediately, and others were monitored, but may have undergone ventilation tube insertion at a later stage, if appropriate.

Four studies enrolled very young children. Maw 1999 randomised children (mean age approximately 3 years) with bilateral OME to receive ventilation tubes or watchful waiting. Paradise 2007 randomised over 400 very young children (mean age 15 months) with either bilateral or unilateral OME to immediate ventilation tubes, or delayed ventilation tube insertion (after a wait of 6 to 9 months). Rovers 2000 randomised young children (mean age approximately 19.5 months) with persistent bilateral OME to insertion of ventilation tubes or watchful waiting. Long term results from the study by Rach 1991 (described above, children aged 2-4) are also included in this comparison, as some participants in the control (no ventilation tube) group underwent ventilation tube insertion during the extended follow-up period.

Two studies considered slightly older children. TARGET 2000 randomised children aged between approximately 3 and 7 years, with bilateral OME, to insertion of ventilation tubes or watchful waiting. A third arm in this trial considered adenoidectomy - data from this arm is relevant for a companion review on the role of adenoidectomy for OME (https://doi.org/10.1002/14651858.CD015252). Velepic 2011 randomised children with predominantly bilateral OME to receive ventilation tube insertion plus adenoidectomy, or adenoidectomy alone.

The child was the unit of analysis for all studies except for Velepic 2011 where the ear was the unit of analysis.

Comparison 3: Ventilation tubes versus non-surgical treatment

A single study was identified for this comparison. Bernard 1991 was a single centre study from Canada, which randomised children to receive either bilateral myringotomy and ventilation tubes, or to receive a 6-month course of antibiotics (sulfisoxazole). Participants were analysed according to their randomised group; however, it should be noted that 47.7% of participants in the medical treatment group did receive ventilation tubes over the course of follow-up, due to 'treatment failure'.

Comparison 4: Ventilation tubes versus myringotomy

We identified 9 studies for this comparison, but different techniques were used to carry out myringotomy.

Laser myringotomy

Two studies randomised children to receive either laser myringotomy or ventilation tubes (D'Eredita 2006; Yousaf 2016). Koopman 2004 enrolled children with bilateral OME, and children received a ventilation tube in one ear and laser myringotomy in the other.

Cold steel myringotomy

Four studies randomised children to receive either bilateral ventilation tubes or cold-steel myringotomy (Gates 1989; Popova 2010; Sujatha 2015; Tao 2020). In addition, half of the children in Gates 1989 and all the children in Popova 2010 received adenoidectomy. One

further RCT randomised children with bilateral OME to receive a ventilation tube in one ear and cold steel myringotomy in the other (To 1984).

Thermal myringotomy

Ruckley 1988 randomised children with bilateral OME to receive a ventilation tube in one ear and thermal myringotomy in the other ear.

Outcomes

Hearing

Return to normal hearing

As with other reviews in this suite, few studies reported our preferred outcome measure for hearing - the proportion of children in whom hearing returns to normal. This outcome was only measured by three studies (D'Eredita 2006; Dempster 1993; Paradise 2007). Dempster 1993 and Paradise 2007 defined 'normal hearing' as <15dBHL, whilst D'Eredita 2006 did not provide a definition.

Final hearing thresholds or change in hearing threshold

The majority of studies assess hearing using mean final hearing thresholds. We have concerns about whether this is an appropriate method to assess hearing, as it may give misleading results - particularly in a condition where there is a high spontaneous resolution. A small mean change in hearing may actually be consistent with a large improvement in hearing for a subset of children (and little change for those who had spontaneous improvement).

Most studies assessed mean hearing thresholds using pure tone audiometry, typically over a range of frequencies (Bernard 1991; Dempster 1993; Maw 1983; Maw 1999; Paradise 2007; Popova 2010; TARGET 2000; To 1984). Rovers 2000 assessed hearing using a portable visual reinforcement audiometry set, which measured the minimal response level (not a mean hearing level) in the better hearing ear. Three studies assessed the air-bone gap when assessing hearing (Ruckley 1988; Sujatha 2015; Velepic 2011).

Disease-specific quality of life

We did not identify any studies which assessed disease-specific quality of life.

Persistent perforation

A small number of studies provided some information about the rate of persistent tympanic membrane perforation.

Persistence of OME

Persistence of OME was assessed in the majority of studies. However, the methods used to identify persistent OME varied - using different combinations of tympanometry, otoscopy and audiometry findings. This may result in some heterogeneity in the effect estimates seen.

Adverse effects: tympanic membrane changes, tube-related, patient-related

Few studies appeared to systematically assess and report on the presence of adverse effects. The data obtained were often not suitable for meta-analysis, as we had insufficient information on the number of events or denominators, or outcomes were only reported for one group.

Receptive language skills

Four studies conducted some kind of assessment of receptive language skills (Maw 1999; Paradise 2007; Rach 1991; Rovers 2000). This outcome was assessed using the Reynell test, the WOLD test, reading fluency Woodcock Reading Mastery Tests, Woodcock-Johnson III Tests of Achievement and tests of phonological processing.

Expressive language skills

The same four studies also assessed expressive language skills, using the Reynell test, WOLD and Schlichting test scores (Maw 1999; Paradise 2007; Rach 1991; Rovers 2000).

Cognitive development

This outcome was assessed by Maw 1999 (using the Griffiths practical reasoning test and the WISC-III short form) and Paradise 2007 (with the Wechsler Abbreviated Scale of Intelligence, and the calculation subset of the Woodcock-Johnson III Tests of Achievement).

Psychosocial outcomes

The study by Maw 1999 considered a number of behavioural outcomes, assessed with the Richman Behaviour Checklist, which is completed by the child's parents (range 0-24, higher scores represent worse behaviour, and a threshold of ≥10 has been suggested as a cut-off to determine behavioural problems). Rovers 2000 used the Erikson Scale of Parent-Child interaction and Paradise 2007 used the Disruptive Behaviour Disorders Rating Scale and Child Behaviour Checklist to assess this outcome.

Listening skills

This outcome was not assessed by any of the included studies.

Generic health-related quality of life

A single study included an assessment of generic health-related quality of life, using the TAIQOL questionnaire (Rovers 2000).

Parental stress

A single study measured this outcome, using the Parenting Stress Index (Paradise 2007).

Vestibular function

This outcome was not assessed by any of the included studies.

Doctor-diagnosed acute otitis media episodes

This outcome was assessed by only two studies (Bernard 1991; Popova 2010).

Excluded studies

We excluded 50 records (linked to 47 studies). The main reasons for exclusion are listed below.

- Eighteen studies were not randomised controlled trials, or did not analyse participants according to their randomised groups (Ah-Tye 2001; Bozkurt 2004; Englender 1999; Ferrara 2005; Gibson 1996; Hassmann 2004; Iino 1989; Kremer 1979; Liu 2004; MRC Multicentre Otitis Media Study 2004; MRC Multicentre Otitis Media Study 2008; Paradise 1997; Parlea 2012; Sanyaolu 2020; Shubich 1996; Stenstrom 2005; Uvarova 2001; Youssef 2013
- Fifteen studies recruited an incorrect population, including:
    - 11 studies in which the duration of OME was unknown, or was definitely less than three months (Black 1990; El Begermy 2022; Bulman 1984; Hammaren-Malmi 2005; Lildholdt 1983; Mandel 1989; Markou 2004; NCT00629694; Rohail 2006; Shishegar 2007; Skinner 1988)
    - three studies in which participants had recurrent acute otitis media, not OME (Gebhart 1981; Kujala 2012; Paradise 1990);
    - one study where participants had acute otitis media (Nguyen 2004).
- Twelve studies where an intervention other than ventilation tubes was assessed. Some of these studies were relevant for other reviews in this suite (Ardehali 2008; Choung 2008; Hao 2019; Jabeen 2019; Mandel 1992; Marchisio 1998; Maw 1993; Moller 1990; NCT05545345; Tao 2020; Xu 2016; Yousaf 2014).

- One study with an incorrect comparator, where ventilation tubes were compared to balloon dilatation of the Eustachian tube (Li 2020).
- One study that was terminated/withdrawn before any results were available (Demant 2017).

Risk of bias in included studies

We had concerns over the potential for bias in all the included studies in this review. See Figure 3 for a summary of the risk of bias across the studies, and Figure 4 for detailed judgements on individual studies.

Allocation

Most studies provided sufficient information regarding the randomisation procedure for us to be confident that a random method was employed. However, seven studies did not provide this information (Bernard 1991; D'Eredita 2006; Dempster 1993; Popova 2010; To 1984; Velepic 2011; Yousaf 2016). One study used quasi-randomisation, where participants were allocated to groups according to the order of recruitment to the study (Elkholy 2021), leading to a high risk of selection bias. Only five studies provided a description of methods used to conceal group allocation (Dempster 1993; Gates 1989; Maw 1999; Ruckley 1988; TARGET 2000). We judged the remaining studies at unclear risk of selection bias, as insufficient information was available to determine whether group allocation may have been predicted

Blinding

None of the included studies appeared to blind participants and study personnel to the intervention received, and only three studies described blinding of outcome assessors (Maw 1999; Paradise 2007; TARGET 2000).

Incomplete outcome data

The risk of bias was mixed for this domain. We considered nine studies to provide sufficient follow-up data that attrition bias was not a concern (Bernard 1991; D'Eredita 2006; Elkholy 2021; Paradise 2007; Sujatha 2015; Tao 2020; To 1984; Velepic 2011; Yousaf 2016). We rated five studies at high risk of attrition bias, due to the level of dropout over the course of the study (Gates 1989; Koopman 2004; Maw 1999; Popova 2010; Rovers 2000). For the remaining studies, there was either insufficient information to judge whether dropout posed a risk of attrition bias, or we were uncertain whether the extent of dropout would be enough to cause a risk here.

Selective reporting

We considered five studies to be at risk of selective reporting, mainly due to incomplete reporting of primary outcome measures (D'Eredita 2006; Koopman 2004; Ruckley 1988; Yousaf 2016). We also rated the study Velepic 2011 at high risk, as it was unclear whether outcome data were provided for follow-up at three months or six months, and raw data are not reported for some outcomes (only P values). The time of follow-up affects interpretation of the outcomes as ventilation tubes were inserted for all participants in the control group who did not have resolution of the effusion after three months.

We rated most of the remaining studies at unclear risk of bias, as no registered protocol was available with which to compare the published reports.

Other potential sources of bias

We identified some additional issues with several studies, which we considered to be a potential risk of bias:

Bernard 1991 used two different types of ventilation tubes over the course of the study, and reported that one was better than the other at improving hearing loss. Data are not available for the different types of ventilation tubes. In addition, many children (48%) in

the control (antibiotics) group also received a ventilation tube over the course of the trial, which may bias the findings towards the null.

Elkholy 2021 only provided useable outcome data after two weeks of follow-up, which is too short to assess the effect of ventilation tubes and no intervention for many outcomes.

Gates 1989 permitted parents to choose a different treatment to the one randomised. This occurred for 5.5% of participants. In addition, many children undergoing medical (49%) or surgical (22%) treatment underwent a second course of the same treatment during the trial.

Popova 2010 appeared to use a 'per protocol' analysis, rather than 'intention-to-treat'.

Ruckley 1988 conducted follow-up at three months, which may be too short to adequately assess the effect of the intervention.

TARGET 2000 retrospectively published the trial protocol, raising the possibility of publication bias. In addition, this was an MRC-funded, multicentre trial and yet not all outcomes stated in the trial registration were published.

To 1984 indicated that most, but not all, children in the control group received a myringotomy. Ideally data would have been available separately for these groups, to include in the comparison of ventilation tubes versus no treatment and ventilation tubes versus myringotomy. The mixed control group may bias the results, if the effect sizes for ventilation tubes versus myringotomy and no treatment differ.

Velepic 2011 only recruited children who regularly attended check-ups, which may have led to a risk of selection bias.

Yousaf 2016 randomised participants at the level of the child, but reported results at the level of the individual ear. This fails to account for correlation between ears of the same individual, and may lead to confidence intervals that are too precise.

Effects of interventions

Comparison 1: Ventilation tubes versus no treatment

Four studies were included in this comparison (Dempster 1993; Elkholy 2021; Maw 1983; Rach 1991).

Hearing

Return to normal hearing at 3 to 12 months follow-up

One study compared the proportion of ears in which hearing returned to normal levels (defined as < 15 dB HL) at 12 months follow-up. The odds ratio (OR) for return to normal hearing was 1.13 in favour of ears which had received ventilation tubes (95% confidence interval (CI) 0.46 to 2.74; 54% versus 51%; 1 study; 72 participants; Analysis 1.1; very low-certainty evidence).

As there is likely to be some correlation in this outcome between ears of the same individual, we attempted to account for this in the analysis. The main analysis was conducted assuming a correlation coefficient of 0.5 between ears of the same individual. However, we conducted sensitivity analyses to determine where changing the assumed correlation would have a significant impact on the results, and it did not (Analysis 5.1; Analysis 5.2).

We also noted that the threshold for 'normal hearing' of < 15 dB HL was lower than we had pre-specified in our protocol. The authors of Dempster 1993 also reported the proportion of ears in which hearing returned to < 25 dB HL. If this threshold was used a 'normal hearing' then there was no difference between the groups, with an OR of 1.00 for ears which receive a ventilation tube (Analysis 5.3).

Final hearing threshold at 3 to 12 months follow-up

Two studies compared the final hearing threshold for ears which had received a ventilation tube, compared to ears which had not, at 12 months follow-up. The mean

difference in hearing level was -3.47 dB HL lower (better) for ears which had received a ventilation tube (95% CI -9.97 to 3.03; 2 studies; 129 participants; Analysis 1.2; very low-certainty evidence).

As above, when we accounted for correlation between the ears of the same individual using a variety of correlation coefficients, the effect size seen was very similar (Analysis 5.4; Analysis 5.5).

Change in hearing threshold at 3 to 12 months follow-up

A single study assessed this outcome at 12 months follow-up. The mean change in hearing level was -0.16 dB HL lower (better) for those ears which received a ventilation tube, compared to those which did not (95% CI -3.28 to 2.97; 1 study; 72 participants; Analysis 1.3; very low-certainty evidence).

Accounting for correlation between ears of the same individual made a very modest difference to the effect estimate, ranging from -0.10 to -0.21 dB HL lower (Analysis 5.6; Analysis 5.7).

Persistent perforation

One study reported on perforation *or retraction* of the tympanic membrane (Dempster 1993). The odds ratio for perforation/retraction was 0.85 for those ears which had received a ventilation tube, compared to those which did not (95% CI 0.38 to 1.91; 8.3% versus 9.7%, 1 study; 72 participants; Analysis 1.4; very low-certainty evidence).

As above, when we accounted for correlation between the ears of the same individual using a variety of correlation coefficients, the effect size seen was very similar (Analysis 5.8; Analysis 5.9).

Persistence of OME

Three studies assessed this outcome. The unit of analysis was different for these trials (Rach 1991 and Elkholy 2021 analysed per child, Dempster 1993 analysed per ear) therefore we have presented the results separately.

Randomised per child

< 6 weeks follow-up

The risk ratio for persistence of OME after just two weeks of follow-up was 0.33 (95% CI 0.08 to 1.46; 10% versus 30%, 1 study; 40 participants; Analysis 1.5; very low-certainty evidence).

3 to 12 months follow-up

After six months, one study reported a risk ratio of 0.30 for persistence of OME in ears that had received ventilation tubes (95% CI 0.14 to 0.65; 20% versus 68%, 1 study; 40 participants; Analysis 1.6; low-certainty evidence). Although the trial was randomised by child, the unit of analysis was the individual ear. Using different intracluster correlation coefficients as part of a sensitivity analysis had little impact on the overall result (Analysis 5.10; Analysis 5.11).

Randomised per ear

One study identified an odds ratio of 0.66 for the persistence of OME in ears that had received ventilation tubes, compared to ears of the same individual that did not have a ventilation tube fitted (95% CI 0.24 to 1.85; 49% versus 58%, 1 study; 72 participants; Analysis 1.7; very low-certainty evidence). We note considerable heterogeneity in the effect between the two different subgroups of children included in this study. The effect size was substantial for those who did not receive adenoidectomy (OR 0.39, 95% CI 0.20 to 0.77), but was trivial for those who did receive adenoidectomy (OR 1.11, 95% CI 0.58 to 2.12).

As above, when we accounted for correlation between the ears of the same individual using a variety of correlation coefficients, the effect size seen was very similar (Analysis 5.12; Analysis 5.13).

Comprehensive language skills

A single study assessed this outcome, using the Reynell test. There was a 0.07 greater mean improvement in the Z score for children who had received bilateral ventilation tubes, as compared to those who did not receive ventilation tubes (95% CI -0.26 to 0.4; 1 study; 43 participants; Analysis 1.8; very low-certainty evidence). We have used Cohen's effect size to interpret these scales, where a change of 0.2 represents a small effect, 0.5 a medium effect and 0.8 a large effect.

Expressive language skills

The same study assessed this outcome, also using the Reynell test. There was a 0.12 greater mean improvement in the Z score for children who had received bilateral ventilation tubes, as compared to those who did not receive ventilation tubes (95% CI -0.27 to 0.51; 1 study; 43 participants; Analysis 1.9; very low-certainty evidence).

Other adverse events

Not all the adverse events reported were amenable to meta-analysis. We have therefore summarised a number of adverse events in Table 3 and Table 4. Additional information is shown in Appendix 3.

Tympanic membrane changes

One study reported a Peto OR of 10.09 for tympanosclerosis in ears which had received a ventilation tube, compared to those which had not (95% CI 4.48 to 22.70; 1 study; 72 participants; Analysis 1.10; low-certainty evidence).

Tube-related changes

Rach 1991 found that in the short term (< 3 months) 9/44 (20.5%) ventilation tubes were in situ and in the medium term (six months), 18/44 (40.9%) of the tubes had extruded in the ventilation tube only group (assessed by otoscopy). Maw 1983 reported that some ventilation tubes were reinserted but no data are presented for the number of extrusions/reinsertions. Dempster 1993 reported that, at the 12-month follow-up visit, 31% of ventilation tubes were still functioning.

Patient-related changes

No patient-related adverse events were reported.

Comparison 2: Ventilation tubes compared to watchful waiting

We included six studies in this comparison. All randomised individual children to receive immediate ventilation tube insertion, or to undergo a period of watchful waiting - with later insertion of ventilation tubes as required.

Hearing

Return to normal hearing

Long-term follow-up (> 1 year)

A single study assessed the proportion of children in whom hearing returned to normal by age 9 to 11, defined as a hearing threshold of ≤ 15 dB HL (Paradise 2007). The risk ratio for return to normal hearing in those with early ventilation tube insertion was 0.98 (95% CI 0.94 to 1.03; 93% compared to 95%, 1 study; 391 participants; Analysis 2.1; very low-certainty evidence).

Mean final hearing threshold

≤ 3 months follow-up

One study assessed final hearing threshold at three months, and found a mean difference of -11.90 dB HL favouring early ventilation tube insertion (95% CI -14.19 to -9.61; 1 study; 215 participants; Analysis 2.2; very low-certainty evidence).

3 to 12 months follow-up

Two studies conducted follow-up at 9 to 12 months. Overall the mean difference in hearing level was -1.89 dB HL in favour of early ventilation tubes (95% CI -7.32 to 3.54; 2 studies; 351 participants; I^2 = 74%; Analysis 2.3; very low-certainty evidence).

One further study also assessed this outcome, but used air-bone gap (rather than air-conduction thresholds). In addition, outcomes were reported per ear (despite randomisation at the level of the individual child) therefore we have had to adjust the results to account for the correlation between ears of the same individual. These results have not been pooled, but show a similar result, with a mean difference of -1.18 dB HL in favour of early ventilation tubes (95% CI -2.9 to 0.54; 1 study; 87 participants with data from 161 ears; Analysis 2.4; very low-certainty evidence). Sensitivity analyses using a different intracluster correlation coefficient showed very similar results (Analysis 6.1; Analysis 6.2).

Long term follow-up (> 1 year)

Three studies conducted follow-up at between 18 months and approximately 3.5 years. The mean difference in hearing threshold for those receiving early ventilation tubes was 0.36 (95% CI -0.41 to 1.13; 3 studies; 633 participants; I^2 = 0%; Analysis 2.5; low-certainty evidence). Sensitivity analyses using a different correlation coefficient for the study Paradise 2007 showed very similar results (Analysis 6.3; Analysis 6.4).

The study Paradise 2007 also assessed hearing using the children's version of the 'hearing in noise' test, where a child repeats sentences heard in a quiet room, and with competing noise. Each sentence is repeated at increasing loudness levels until the child can hear and repeat it. As above, the differences between the two groups were trivial (mean difference ranged from 0 dB to 0.4 dB higher; 1 study; 391 participants; Analysis 2.6; very low-certainty evidence).

Change in hearing threshold from baseline

3 to 12 months follow-up

One study assessed the change in hearing over the course of the study. The mean difference in hearing threshold between the two groups was -4.60 dB HL in favour of early ventilation tubes at between 9 and 12 months of follow-up (95% CI -8.57 to -0.63; 1 study; 176 participants; Analysis 2.7; very low-certainty evidence).

This study also reported a multivariate analysis of the difference in hearing improvement between the two groups, adjusted for baseline hearing level and age. Here the mean difference was -1.6 dB better for those receiving early ventilation tubes (95% CI -0.62 to 3.82; 1 study; 166 participants; Analysis 2.8; very low-certainty evidence).

Adverse event: persistent perforation

3 to 12 months follow-up

One study assessed the rate of persistent tympanic membrane perforations after six months of follow-up, but reported no events in either group (risk difference 0, 95% CI -0.03 to 0.03; 1 study; 161 participants; Analysis 2.9; very low-certainty evidence).

In the TARGET 2000 trial, of 635 ears that had a ventilation tube inserted, eight had a perforation recorded at least six months after surgery. However, of the four who attended later appointments, all had healed.

Long term follow-up (> 1 year)

One study assessed the rate of perforation after approximately 3.5 years of follow-up. The risk ratio for perforation for those who had received early ventilation tubes was 3.65 (95% CI 0.41 to 32.38; 1 study; 281 participants, but data are reported according to ears affected and adjusted for correlation between ears of the same individual; Analysis 2.10; very low-certainty evidence).

Presence/persistence of OME

3 to 12 months follow-up

Three studies assessed this outcome, but used slightly different ways of assessing and reporting persistent OME. Velepic 2011 assessed persistence of OME in both ears using otoscopy at six months follow-up, and found a risk ratio of 0.39 for participants who had undergone early ventilation tube insertion (95% CI 0.09 to 1.70; 5% versus 13%, 1 study; 87 participants; Analysis 2.11; very low-certainty evidence).

Maw 1999 used tympanometry to assess the presence of OME in the better ear at nine months of follow-up and found a risk ratio of 0.52 for those who had undergone early ventilation tube insertion (95% CI 0.37 to 0.71; 37% versus 70%, 154 participants; Analysis 2.12; low-certainty evidence). Finally, Paradise 2007 reported on the percentage of days during follow-up that OME persisted for in each group. OME persisted for 19% fewer days in those who had received early ventilation tubes (95% CI 23% fewer to 15% fewer; 1 study; 316 participants; Analysis 2.13; very low-certainty evidence).

Long term follow-up (> 1 year)

Three studies assessed the presence or persistence of OME after long-term follow-up using tympanometry (from 18 months to over six years) and found a risk ratio of 1.21 for those who had undergone early ventilation tube insertion (95% CI 0.84 to 1.74; 15% versus 12%, 3 studies; 584 participants; I^2 = 0%; Analysis 2.14; very low-certainty evidence).

One of these studies also presented an adjusted effect estimate, accounting for baseline differences in gender, age, housing status, maternal education and mother's parity. The odds ratio for abnormal tympanometry was 0.99 (95% CI 0.35 to 2.83; 1 study; 65 participants; Analysis 2.15; very low-certainty evidence).

Adverse events

Adverse events were reported inconsistently by the different studies, and many were not amendable to analysis. We have therefore summarised a number of adverse events in Table 3 and Table 4. Additional information is shown in Appendix 3.

Receptive language skills

Three studies assessed receptive language skills at medium term (Maw 1999; Rovers 2000) and long-term follow-up (Maw 1999; Paradise 2007). This outcome was assessed using the Reynell test, the WOLD test, reading fluency Woodcock Reading Mastery Tests, Woodcock-Johnson III Tests of Achievement and tests of phonological processing. Overall, outcomes on these tests either showed a trivial difference between the two groups, or slight benefit for those who received early ventilation tubes (see Analysis 2.20; Analysis 2.21; Analysis 2.22; Analysis 2.23; Analysis 2.24; Analysis 2.25; Analysis 2.26; Analysis 2.50; Analysis 2.51 and Table 5). However, we assessed all the evidence as very low certainty.

Expressive language skills

The same studies also assessed expressive language skills at medium-term (Maw 1999; Rovers 2000) and long-term follow-up (Maw 1999), using the Reynell test, WOLD and Schlichting test scores. Again, the difference between the two groups was largely trivial, or showed a very slight benefit to early ventilation tubes, but the evidence was all very low-certainty (see Analysis 2.27; Analysis 2.28; Analysis 2.29; Analysis 2.30; Analysis 2.31; Analysis 2.32; Analysis 2.33). Some additional data from Paradise 2007 are reported in Table 5.

A number of other aspects of language development were assessed by Maw 1999 after long-term follow-up, including repetition of nonsense words (using the CN/Rep), reading ability (using the WORD test), spelling ability (using 15 age-appropriate words to spell) and an assessment of the ability to delete phonemes when repeating a word (using the Auditory analysis test). Again, the evidence for these outcomes was very low-certainty (see Analysis 2.34; Analysis 2.35; Analysis 2.36; Analysis 2.37).

Cognitive development

Maw 1999 assessed cognitive development at nine months (using the Griffiths practical reasoning test) and 18 months (using the WISC-III short form), but the evidence was very low-certainty (Analysis 2.38; Analysis 2.39). Paradise 2007 also assessed cognition (with the Wechsler Abbreviated Scale of Intelligence, and the calculation subset of the Woodcock-Johnson III Tests of Achievement). No difference was seen between the two groups, but the evidence was very low-certainty. Some additional data from Paradise 2007 are reported in Table 5.

Psychosocial outcomes

The study by Maw 1999 considered a number of behavioural outcomes, assessed with the Richman Behaviour Checklist, which is completed by the child's parents (range 0 to 24, higher scores represent worse behaviour, and a threshold of ≥ 10 has been suggested as a cut-off to determine behavioural problems). At medium-term follow-up, scores were very slightly lower (better) for those who received early ventilation tubes (mean difference -0.65, 95% CI -1.85 to 0.55; 1 study; 150 participants; Analysis 2.40) and the risk ratio for behavioural problems was lower for those receiving early ventilation tubes (RR 0.63, 95% CI 0.42 to 0.96; 1 study; 150 participants; Analysis 2.41). However, the evidence was very low certainty, and adjustment for potential confounding factors (including hearing level) resulted in a change in the direction of the effect. The adjusted odds ratio was 1.16 for behavioural problems in those who received early ventilation tubes, although the confidence intervals were extremely wide (95% CI 0.27 to 4.90; 1 study; 150 participants; Analysis 2.42; very low-certainty evidence).

At longer-term follow-up (18 months) behavioural scores were very slightly worse for those who received early ventilation tubes, but the difference between the groups may be trivial, and the evidence was all very low-certainty (1 study; 123 participants; Analysis 2.43; Analysis 2.44; Analysis 2.45). Similar results were seen by the study from Paradise 2007 when rating behaviour, social skills and continuous performance tests (see Analysis 2.52; Analysis 2.53; Analysis 2.54 and Table 5).

Interaction between parents and children was also assessed in the study Rovers 2000, and a trivial difference was seen in outcomes between the two groups, but the evidence was very low-certainty (see Analysis 2.46; Analysis 2.47).

Parental stress

A single study measured this outcome, using the Parenting Stress Index, but there was no evidence of a difference in parental stress between the two groups after long-term follow-up (mean difference 0, 95% CI -4.12 to 4.12; 1 study; 383 participants; Analysis 2.48; very low-certainty evidence).

Generic health-related quality of life

One study assessed quality of life using the TAIQOL questionnaire (Rovers 2000). A trivial difference was found between the groups across all domains studied, but the evidence was very low-certainty (see Analysis 2.49).

Comparison 3: Ventilation tubes versus non-surgical treatment

This comparison included a single study (Bernard 1991).

Hearing

Final hearing threshold

At short-term follow-up (two months) the mean final hearing threshold was -9 dB HL lower (better) for those who received ventilation tubes, as compared to medical treatment (95% CI -12.61 to -5.39; 1 study; 125 participants; Analysis 3.1; very low-certainty evidence). At medium term follow-up (four months) the mean difference was -5.98 dB HL lower (95% CI -9.21 to -2.75; 1 study; 125 participants; Analysis 3.2; very low-certainty evidence).

Adverse events

The prevalence of most adverse events was only reported for those who had received ventilation tubes. Data on adverse events reported in this study are presented in Table 3 and Table 4, and Appendix 3.

Number of doctor-diagnosed acute otitis media (AOM) episodes

At medium-term follow-up the number of doctor-diagnosed episodes of AOM was lower in those who received ventilation tubes, with a mean difference of -0.23 episodes fewer (95% CI -0.42 to -0.04; 1 study; 125 participants; Analysis 3.4; very low-certainty evidence). However, the difference between the two groups was trivial after long-term follow-up (mean difference -0.05 episodes fewer, 95% CI -0.31 to 0.21; 1 study; 125 participants; Analysis 3.5; very low-certainty evidence).

Comparison 4: Ventilation tubes versus myringotomy

We identified nine studies for this comparison, but they used different techniques to carry out myringotomy (D'Eredita 2006; Gates 1989; Koopman 2004; Popova 2010; Ruckley 1988; Sujatha 2015; Tao 2020; To 1984; Yousaf 2016).

Hearing

Return to normal hearing

Two studies assessed the proportion of participants in whom hearing returned to normal (at six months and one year of follow-up). The risk ratio for return to normal hearing was 1.22 for those who received ventilation tubes compared to laser myringotomy (95% CI 0.59 to 2.53; 74% versus 64%, 2 studies; 120 participants but data reported per ear; I^2 = 95%; Analysis 4.1; very low-certainty evidence). Sensitivity analysis with the use of different intracluster correlation coefficients made very little difference to the overall estimates (see Analysis 7.1; Analysis 7.2).

Final hearing threshold

≤ 3 months follow-up

Two studies assessed this outcome in the short term, but we did not pool the results as one study reported the number of ears affected, and one reported the number of children affected. Both found a trivial difference between the groups in final hearing threshold at short-term follow-up (mean difference for those receiving ventilation tubes 0.2 dB HL higher for one study (95% CI 1.71 to 2.11; 156 participants), and 4.3 dB HL lower for the other study (95% CI -8.55 to -0.05; 108 participants); Analysis 4.2; Analysis 4.3; and see sensitivity analyses Analysis 7.5; Analysis 7.6), but the evidence was very low-certainty.

3 to 12 months follow-up

One study also assessed hearing at 12 months of follow-up and, again, found a trivial difference between the groups, but the evidence was very low-certainty (MD 0.80 dB HL, 95% CI -0.87 to 2.47; 1 study; 156 participants; Analysis 4.4; very low-certainty evidence).

Adverse event: persistent perforation

Only two studies clearly reported the rate of persistent perforation in both groups of participants, allowing a comparison to be made between the groups, however the evidence was all very low-certainty. After three months, Yousaf 2016 identified one perforation in the ears that received laser myringotomy, and two in the ears that received ventilation tubes. Accounting for the potential for correlation between ears of the same individual gave a risk ratio of 1.00 (95% CI 0.06 to 15.50; 1 study; 90 ears; Analysis 4.5; very low-certainty evidence), although if the correlation between ears was less than the risk ratio was higher (see Analysis 7.7; Analysis 7.8).

There appeared to be an increase risk of perforation with ventilation tubes compared with cold-steel myringotomy, but the evidence was very uncertain (Peto OR 8.09, 95% CI 1.78 to 36.79; 1 study; 208 participants; I^2 = 0%; Analysis 4.6; very low-certainty evidence). In

addition, Gates 1989 reported that six children had a persistent perforation of the tympanic membrane: three in the myringotomy group and three who received ventilation tubes. However, the number assessed in each group was not reported, therefore we could not include these data in the meta-analysis.

In the D'Eredita 2006 study, one child in the ventilation tubes group required “myringoplasty to close a persistent TM perforation after 1 year”. No data were reported for the myringotomy group, but it is unclear whether this is because no persistent perforations occurred, or this outcome was not assessed in the group.

Persistence of OME

≤ 3 months follow-up

Two studies assessed the persistence of OME in the short term, but used different types of myringotomy. Yousaf 2016 compared ventilation tubes to laser myringotomy and found a risk ratio of 1.40 for persistence of OME in those receiving ventilation tubes, although the confidence interval was wide (95% CI 0.48 to 4.08; 14% versus 10%, 1 study; 90 participants; Analysis 4.7; very low-certainty evidence). Sensitivity analyses to account for the correlation between ears made little difference to the overall estimates (Analysis 7.9; Analysis 7.10).

Ruckley 1988 compared ventilation tubes with thermal myringotomy. The result was a Peto OR of 0.11 for persistence of OME in those receiving ventilation tubes (95% CI 0.02 to 0.53; 0% versus 19%, 1 study; 72 participants; Analysis 4.8; very low-certainty evidence).

3 to 12 months follow-up

Three studies considered persistence of OME at medium term follow-up. The point estimate for each study showed a benefit to ventilation tubes as compared to myringotomy, however the confidence intervals were very wide and the evidence was all very low-certainty:

- Ventilation tubes versus cold-steel myringotomy: RR 0.69 (95% CI 0.20 to 2.36; 1 study; 78 participants; Analysis 4.9; very low-certainty).
- Ventilation tubes versus laser myringotomy: RR 0.32 (95% CI 0.15 to 0.67; 1 study; 90 participants; Analysis 4.10; very low-certainty evidence). Sensitivity analysis to account for correlation between ears of the same individual made little difference to the overall effect estimates (Analysis 7.11; Analysis 7.12).
- Ventilation tubes versus laser myringotomy, randomised by ear: OR 0.27 (95% CI 0.19 to 0.38; 1 study; 272 ears; Analysis 4.11; very low-certainty evidence). Sensitivity analysis to account for correlation between ears of the same individual made little difference to the overall effect estimates (Analysis 7.13; Analysis 7.14).

One study assessed persistence of OME slightly differently, considering the number of days before the recurrence of OME in each group. Gates 1989 reported a mean difference of 173.88 days longer before recurrence in those who received ventilation tubes as compared to myringotomy (95% CI 150.19 to 197.56; 1 study; 389 participants; Analysis 4.12; very low-certainty evidence).

Long term follow-up (> 1 year)

One study considered persistence of OME in the long term, and found little difference between the two groups after two years of follow-up (RR 0.97, 95% CI 0.90 to 1.05; 83% versus 85%, 1 study; 491 participants; Analysis 4.13; very low-certainty evidence).

Tao 2020 also reported recurrence of OME at 3, 6 and 12 months. However, they also describe additional "conservative treatment" received by these patients. It is not clear what this conservative treatment is, and whether it was balanced across the two groups, so we have not presented these findings.

Episodes of acute otitis media

Only one study assessed the occurrence of acute otitis media during the follow-up period. This was reported as the proportion of participants who experienced a specific number of episodes over the course of 12-month follow-up. The evidence was all very low-certainty.

- No episodes of AOM for those receiving ventilation tubes compared to myringotomy: RR 0.95 (95% CI 0.73 to 1.25; 1 study; 78 participants: Analysis 4.15).
- One episode of AOM for those receiving ventilation tubes compared to myringotomy: RR 1.00 (95% CI 0.37 to 2.71; 1 study; 78 participants; Analysis 4.15).
- Two episodes of AOM for those receiving ventilation tubes compared to myringotomy: RR 0.86 (95% CI 0.18 to 3.99; 1 study; 78 participants; Analysis 4.15).
- Three episodes of AOM for those receiving ventilation tubes compared to myringotomy: Peto OR 6.41 (95% CI 0.13 to 326.59; 1 study; 78 participants; Analysis 4.16).
- Four or more episodes of AOM for those receiving ventilation tubes compared to myringotomy: Peto OR 6.41 (95% CI 0.13 to 326.59; 1 study; 78 participants; Analysis 4.16).

Adverse events

Details are reported in Appendix 3, Table 3 and Table 4.

Discussion

There are some certainties in otitis media with effusion (OME). Firstly, this is a fluctuating condition with a high rate of spontaneous resolution, but also a high rate of recurrence over time. The impact of OME on any individual child is very variable, and consequently the need for treatment differs. So far, attempts to understand the condition better with prognostic studies have been unsuccessful.

In undertaking this review, and using the GRADE approach to assess the certainty of evidence (according to Cochrane methodology), we have encountered a high degree of 'uncertainty'. The GRADE approach considers methodological rigour of the studies, but also looks at precision of the effect estimates, applicability of the results and inconsistency in estimates between different studies. Despite the large number of studies included in the review, limited pooling of data was possible. Relatively small numbers of participants were included in many analyses, resulting in wide confidence intervals for measures of effect.

There are still key questions that remain unanswered in this common disease. Resolving these uncertainties is absolutely critical to enable research in this area to progress.

Firstly, we need to identify which children will undergo spontaneous resolution of OME, through a better understanding of prognostic factors in the disease. This would allow treatments to be targeted to those children in whom OME is more likely to be persistent, and impact language and development. Many of the studies included in this review recruited a variety of children - some with unilateral OME, and some with mild hearing loss. It is possible that these children are less likely to benefit from any intervention to treat OME, as the disease may have little impact on their development and quality of life. Including these children in trials may result in an under-estimate of the efficacy of the intervention, and bias the overall results towards the null.

In addition, although our primary outcome measure was hearing, we are aware that this is not the only important factor in this disease. Children with identical levels of hearing loss from OME may have very different outcomes in terms of the impact of the disease on development and quality of life. Again, a clearer understanding of the disease process, and different subgroups of children with OME would help to identify those children who are at risk of poor long-term outcomes.

Summary of main results

All the evidence identified in this review was either low- or very low-certainty, showing that we have little confidence in the overall estimates of effect.

Ventilation tubes compared to no treatment

There were very few trials that assessed this comparison, as it does not reflect routine clinical practice, where patients would be offered either immediate surgery, or a period of watchful waiting. After 12 months, there appeared to be little difference in the proportion of children whose hearing returned to normal with or without ventilation tubes. The mean difference in hearing threshold was also small, although we have concerns about the use of mean hearing thresholds to assess hearing in this context (see below). Overall, persistence of OME appeared slightly lower for those who received ventilation tubes (at follow-up of up to one year). Little difference was seen between the two groups for receptive and expressive language skills. Very limited data on adverse events were available.

Ventilation tubes compared to watchful waiting

After long-term follow-up there was little difference in the proportion of children whose hearing had returned to normal. When final hearing threshold was assessed, there was a benefit to ventilation tubes at short-term follow-up (three months), but this reduced after longer-term follow-up. This may be due to the high proportion of children in the control group who underwent surgery during the follow-up period. Persistence of OME appeared to be reduced after six to nine months for those who received ventilation tubes, but this effect was not seen after longer-term follow-up. Very limited data on adverse events were available. Evidence for expressive language skills, receptive language skills, cognitive development, psychosocial outcomes, parental stress and generic quality of life was all very low-certainty, but little difference was seen between the two groups.

Ventilation tubes versus non-surgical treatment

A single study compared ventilation tubes to long-term antibiotic treatment. The mean final hearing threshold was slightly better for those who received ventilation tubes, but very few data were reported for other outcomes.

Ventilation tubes compared to myringotomy

There may be a slight increase in the proportion of children whose hearing returns to normal with ventilation tubes (as compared to myringotomy). Very little difference in the mean final hearing threshold was seen but, as described below, we are uncertain if this method of assessing hearing is appropriate for this condition. The rate of persistent tympanic membrane perforation is probably increased with ventilation tubes as compared to myringotomy. After medium-term follow-up, ventilation tubes may slightly reduce the rate of persistent OME. However, this effect was not seen after longer-term follow-up. Very limited data on adverse events were available.

Overall completeness and applicability of evidence

The focus of this review was to summarise the evidence from randomised controlled trials (RCTs). However, in a condition such as OME - with very variable effects on individual children, fluctuating symptoms and little understanding of important prognostic factors - an RCT may not be the preferred study design. The review does not include data from large cohort studies, which have highlighted the fluctuation of symptoms of OME in those both with and without ventilation tubes (Zielhuis 1990).

In keeping with other reviews in this suite, we noted that very few studies reported our preferred outcome measure for hearing - the number of children who returned to normal hearing. We have concerns that assessment of hearing using the mean difference in final hearing threshold (or mean change in hearing threshold) may not be the most appropriate

way to assess hearing. OME has a high spontaneous resolution rate. Consequently, we would anticipate that the change in hearing threshold for most children will be similar across the groups, as many children will improve with or without treatment. Therefore, even if a subset of children had substantial benefit from the intervention, the overall mean difference between the two groups would appear to be small. When assessed using the mean difference, the marked benefit seen in a subgroup of participants is ‘diluted’ by the children who get better regardless of treatment. Therefore, an apparently small mean difference between the two groups may actually be consistent with a substantial change in the number of children in whom hearing returns to normal.

Interpreting the results of the comparison between ventilation tube insertion and watchful waiting is challenging. This situation is commonly encountered in clinical practice, where children, their parents and healthcare professionals may need to decide between immediate insertion of ventilation tubes or a further period of watchful waiting. However, the high rate of ventilation tube insertion in the watchful waiting group means that it is difficult to understand the effect of ventilation tubes. The similarities between the intervention and control groups after long-term follow-up may be because of spontaneous improvement in symptoms, but also may be because of the high rate of intervention in the control group. In addition, ventilation tubes become blocked, and will extrude over time, and OME can recur. Comparing the prevalence of OME in those who received and did not receive ventilation tubes therefore becomes more difficult to interpret after longer-term follow-up.

The results of this review should be assessed in conjunction with those of the companion review regarding the use of adenoidectomy for OME (MacKeith 2022a). It is possible that there are synergistic effects of ventilation tubes and adenoidectomy when treating OME. Many of the studies included in this review provided adenoidectomy as a background intervention to all children. The effect of ventilation tubes on OME may be modified in children who also receive adenoidectomy. For example, if children receiving adenoidectomy already have a high rate of resolution for OME, then any additional benefit of ventilation tubes may not be clearly identified.

Quality of the evidence

We considered most of the evidence included in this review to be very low-certainty. This was predominantly due to concerns over the risk of bias in the studies included, particularly the risk of performance and detection bias. However, many studies also had unclear ratings for the risk of selection bias, attrition bias or reporting bias. In addition, many of the studies included relatively few participants, which led to wide confidence intervals and imprecision in the overall effect estimates.

Potential biases in the review process

We have attempted to minimise the potential for bias during the review process by adhering to the *Cochrane Handbook for Systematic Reviews of Interventions* throughout the conduct of this review. We conducted comprehensive searches and ensured that study selection, data extraction and GRADE assessment were carried out by at least two independent authors, to ensure reproducibility of findings.

Agreements and disagreements with other studies or reviews

The results of this review are similar to that from the previous Cochrane Review on this topic, which included 10 studies (Browning 2010). At that time, the authors concluded that the effects of ventilation tubes on hearing appears to be small, and reduces after six to nine months. The time with effusion (analogous to our outcome 'persistence of OME') was reduced for those who received ventilation tubes. Again, this benefit was smaller after longer follow-up.

In accordance with current Cochrane standards we have now used the GRADE approach to assess the certainty of the evidence; the previous Cochrane Review on this topic pre-

dated the GRADE criteria. This approach means that our conclusions appear less certain than the previous review, but it should be noted that the evidence has not changed, it is simply that we are looking at the data with a new approach.

Authors' conclusions

Implications for practice

Whilst there may be small short-term improvements in hearing and persistence of otitis media with effusion (OME) with ventilation tubes, it is unclear whether there are lasting benefits when children are followed up for longer periods of time. There is a risk of complications from surgery, including persistent tympanic membrane perforation. The extent of this risk is unclear, but is likely to be small.

Most of the studies in this review specifically excluded children with risk factors for OME, such as cleft palate or Down syndrome. Therefore, we do not have any information on the efficacy or harms of this intervention for those children. We were also unable to carry out our planned subgroup analyses, to determine if the effect of ventilation tubes may vary across children of different ages, different levels of hearing loss or with co-morbidities.

Implications for research

This review forms part of a suite of five reviews that consider interventions for otitis media with effusion (OME) (Galbraith 2022; MacKeith 2022a; MacKeith 2022b; Mulvaney 2022a; Mulvaney 2022b). Here we present implications for research in this field, which are shared across the suite of reviews:

1. As OME is a fluctuating condition with high rates of resolution and recurrence, and a highly variable impact on children, clinical trials (and, in particular, randomised controlled trials) may not be the research design of choice. Instead, evidence may be better obtained from surgical or clinical registries (for example, see Schmalbach 2021) or prospective cohort studies, with the use of 'big data'. These data sets may also be used to help identify subgroups of children who are at greater risk of persistent disease or long-term consequences of OME. A clearer understanding of possible subgroups of children is needed to better target interventions to those who need them most, whilst avoiding over-treatment for those in whom spontaneous resolution is anticipated.
2. Adverse effects of interventions are important, and should always be assessed. However, randomised controlled trials are also not the best method to consider these, especially when events are rare. Observational studies with longer follow-up and larger numbers of participants are needed to provide more robust evidence on the frequency of side effects.
3. It is encouraging that a core outcome set has been developed in this field (Bruce 2015; Liu 2020). Guidance on *how* to measure the different outcomes would also be helpful for future research.
4. Comparison of mean hearing thresholds is widely used in research to assess the impact of different interventions on hearing. However, this outcome measure risks underestimating the potential impact of interventions on hearing. Small changes in mean hearing thresholds may be consistent with a substantial improvement in the number of children whose hearing returns to normal, particularly in a condition with a high spontaneous resolution rate. We would encourage researchers to assess hearing with the proportion of children in whom hearing returns to normal, in preference to mean hearing thresholds.

Acknowledgements

This project was supported by the National Institute for Health Research, via Cochrane Infrastructure, Cochrane Programme Grant or Cochrane Incentive funding to Cochrane ENT. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Systematic Reviews Programme, NIHR, NHS or the Department of Health.

The authors are grateful to Anne Littlewood, Information Specialist with Cochrane Oral Health, for providing peer review comments on the draft search methods.

We are also grateful to Dr Richard Rosenfeld for providing editorial sign-off of the protocol. *[To be updated after peer review/sign-off]*

We are grateful to the 20 Cochrane Crowd screeners for screening 2443 records to identify 1130 possible RCTs, and reject 1313 references as not RCTs. We are particularly grateful to Bernardo Costa, Stefanie Rosumeck, Nikolaos Sideris, Susanna Wisniewski, Anna Resolver, Lai Ogunsola, Shammas Mohammed, Sarah Moore, Brian Duncan, Mohammad Aloulou, Ana-Marija Ljubenković, Vighnesh D, Ahlam Jamal Alhemedi, Neetu Bhadra, Amin Sharifan, Abu Emmil Qawarizmi Bin Abu Sofian, Helen Ramsay, Dinah Amoah, Maike Scherhans and Natalya Clark for screening more than 200 records each.

Finally, our grateful thanks to Jenny Bellorini, Managing Editor for Cochrane ENT, and Samantha Cox, Information Specialist, without whom the development of these reviews would not have been possible.

Editorial and peer reviewer contributions

[To be completed after peer review/sign-off] Cochrane ENT supported the authors in the development of this review.

The following people conducted the editorial process for this article:

- Sign-off Editor (final editorial decision): [NAME, AFFILIATION];
- Managing Editor (selected peer reviewers, collated peer reviewer comments, provided editorial guidance to authors, edited the article): [NAME, AFFILIATION];
- Copy Editor (copy editing and production): [NAME, AFFILIATION];
- Peer reviewers (provided comments and recommended an editorial decision): [NAME, AFFILIATION] (clinical/content review)*, [NAME, AFFILIATION] (consumer review), [NAME, AFFILIATION] (methods review), [NAME, AFFILIATION] (search review). [NUMBER] of additional peer reviewers provided [CLINICAL/CONTENT/CONSUMER/METHODS/SEARCH] peer review, but chose not to be publicly acknowledged.

Data and analyses

Comparison 1

VT versus no treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
1.1 Return to normal hearing, randomised by ear (medium-term)	1		Odds Ratio (IV, Random, 95% CI)	1.13 [0.46, 2.74]
1.1.1 Randomised by ear: normal defined as <15dB. CC=0.5 (medium term)	1		Odds Ratio (IV, Random, 95% CI)	1.13 [0.46, 2.74]
1.2 Mean final hearing threshold, randomised by ear (medium-term)	2		Mean Difference (IV, Random, 95% CI)	-3.47 [-9.97, 3.03]
1.2.1 Correlation coefficient = 0.5	2		Mean Difference (IV, Random, 95% CI)	-3.47 [-9.97, 3.03]
1.3 Change in hearing threshold	1		Mean Difference	-0.16 [-3.28, 2.97]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size	
from baseline, randomised by ear (medium-term)			(IV, Random, 95% CI)		
1.4 Adverse event: perforation/retraction, randomised by ear (medium-term)	1		Odds Ratio (IV, Random, 95% CI)	Subtotals only	
1.4.1 Correlation coefficient 0.5	1		Odds Ratio (IV, Random, 95% CI)	0.85 [0.38, 1.91]	
1.5 Persistence of OME: randomised by child (very short-term)	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected	
1.6 Persistence of OME: randomised by child (medium-term)	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected	
1.6.1 Adjusted for non-independence of within-individual measurements, assuming ICC of 0.5	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected	
1.7 Persistence of OME: randomised by ear (medium-term)	1		Odds Ratio (IV, Random, 95% CI)	Subtotals only	
1.7.1 Correlation coefficient = 0.5	1		Odds Ratio (IV, Random, 95% CI)	0.66 [0.24, 1.85]	
1.8 Mean improvement in comprehensive language, randomised by child (medium-term)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only	
1.9 Mean improvement in expressive language, randomised by child (medium-term)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only	
1.10 Adverse event: tympanosclerosis, randomised by ear (medium-term)	1	144	Peto Odds Ratio (Peto, Fixed, 95% CI)	10.09 [4.48, 22.70]	

Comparison 2

Early VT versus watchful waiting (treatment later if required)

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size	
2.1 Hearing returned to normal, randomised by child (long-term)	1	391	Risk Ratio (M-H, Random, 95% CI)	0.98 [0.94, 1.03]	
2.2 Mean final hearing threshold, randomised by child (short-term)	1		Mean Difference (IV,	Subtotals only	

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
			Random, 95% CI)	
2.3 Mean final hearing threshold (air conduction), randomised by child (medium-term)	2	351	Mean Difference (IV, Random, 95% CI)	-1.89 [-7.32, 3.54]
2.4 Mean final hearing threshold (air-bone gap), randomised by child, analysed by ear (medium-term)	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
2.4.1 Adjusted for non-independence of within-individual measurements, assuming ICC of 0.5	1		Mean Difference (IV, Random, 95% CI)	Totals not selected
2.5 Mean final hearing threshold, randomised by child (long-term)	3		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.5.1 Assume correlation coefficient for Paradise 2007 (left and right ear data combined) of 0.5	3	633	Mean Difference (IV, Random, 95% CI)	0.36 [-0.41, 1.13]
2.6 Hearing in noise test, randomised by child (long-term)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.6.1 Competing noise from the front (dB)	1	391	Mean Difference (IV, Random, 95% CI)	0.20 [-0.13, 0.53]
2.6.2 Competing noise from the right (dB)	1	391	Mean Difference (IV, Random, 95% CI)	0.00 [-0.54, 0.54]
2.6.3 Competing noise from the left (dB)	1	391	Mean Difference (IV, Random, 95% CI)	0.40 [-0.10, 0.90]
2.7 Change in hearing threshold from baseline, randomised by child (medium-term)	1	176	Mean Difference (IV, Random, 95% CI)	-4.60 [-8.57, -0.63]
2.8 Adjusted mean difference in hearing improvement, randomised by child (medium term)	1		Mean Difference (IV, Random, 95% CI)	1.60 [-0.62, 3.82]
2.9 Adverse event: persistent perforation, randomised by child (medium-term)	1	161	Risk Difference (M-H, Random, 95% CI)	0.00 [-0.03, 0.03]
2.10 Adverse event: persistent perforation,	1		Risk Ratio (M-H,	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
randomised by child (long-term)			Random, 95% CI)	
2.10.1 Adjusted for non-independence of within-individual measurements: ICC 0.5	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
2.11 Presence/persistence of OME, randomised by child, measured by otoscopy (medium-term)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
2.11.1 Adjusted for non-independence of within-individual measurements, assuming ICC of 0.5	1	113	Risk Ratio (M-H, Random, 95% CI)	0.39 [0.09, 1.72]
2.12 Presence/persistence of OME, randomised by child, measured by tympanometry (medium-term)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
2.13 Presence/persistence of OME, mean percentage of days, randomised by child (medium-term)	1	316	Mean Difference (IV, Random, 95% CI)	-0.19 [-0.23, -0.15]
2.14 Presence/persistence of OME, randomised by child (long-term)	3	584	Risk Ratio (M-H, Random, 95% CI)	1.21 [0.84, 1.74]
2.15 Presence/persistence of OME, adjusted OR, randomised by child (long-term)	1		Odds Ratio (IV, Random, 95% CI)	0.99 [0.35, 2.83]
2.16 Adverse event: tympanosclerosis (long term)	1	375	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.33, 2.55]
2.16.1 Adjusted for non-independence of within-individual measurements: ICC 0.5	1	375	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.33, 2.55]
2.17 Adverse event: fibrosis (long term)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
2.17.1 Adjusted for non-independence of within-individual measurements: ICC 0.5	1	375	Risk Ratio (M-H, Random, 95% CI)	0.61 [0.10, 3.60]
2.18 Adverse event: segmental atrophy (long term)	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
2.18.1 Adjusted for non-independence of	1		Risk Ratio (M-	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
within-individual measurements. Assumed ICC 0.5			H, Random, 95% CI)	
2.19 Adverse event: retraction pocket with other abnormality (long term)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
2.19.1 Adjusted for non-independence of within-individual measurements. ICC assumed 0.5	1	374	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.06, 14.41]
2.20 Receptive language development, Reynell test, randomised by child (medium-term)	1		Mean Difference (IV, Random, 95% CI)	0.31 [-0.03, 0.65]
2.21 Receptive language development, Reynell test, adj MD (medium-term)	1		Mean Difference (IV, Random, 95% CI)	0.39 [0.04, 0.74]
2.22 Receptive language, Reynell test, randomised by child (long-term)	1		Mean Difference (IV, Random, 95% CI)	0.26 [-0.08, 0.60]
2.23 Receptive language: Reynell test, long-term, adjusted MD	1		Mean Difference (IV, Random, 95% CI)	0.17 [-0.21, 0.55]
2.24 Receptive language: WOLD adjusted OR (long-term)	1		Odds Ratio (IV, Random, 95% CI)	1.58 [0.59, 4.24]
2.25 Receptive language, mean difference (months) in improvement in Reynell test score (equivalent age -real age): medium-term	1		Mean Difference (IV, Random, 95% CI)	1.01 [-0.14, 2.16]
2.26 Receptive language, adjusted mean difference (months) in improvement in Reynell test score (equivalent age - real age): medium-term	1		Mean Difference (IV, Random, 95% CI)	0.71 [-0.28, 1.70]
2.27 Expressive language development: Reynell test (medium-term)	1		Mean Difference (IV, Random, 95% CI)	0.38 [-0.00, 0.76]
2.28 Expressive language development: Reynell test, medium-term, adjusted MD	1		Mean Difference (IV, Random, 95% CI)	0.42 [0.02, 0.82]
2.29 Expressive language development: Reynell test (long-term)	1		Mean Difference (IV, Random, 95% CI)	0.31 [-0.07, 0.69]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.30 Expressive language development: Reynell test, long-term, adjusted MD	1		Mean Difference (IV, Random, 95% CI)	0.14 [-0.28, 0.56]
2.31 Expressive language: WOLD adjusted OR (long-term)	1		Odds Ratio (IV, Random, 95% CI)	2.10 [0.78, 5.65]
2.32 Expressive language, mean difference (months) in improvement in Schlichting test score (equivalent age -real age): medium-term	1		Mean Difference (IV, Random, 95% CI)	-0.53 [-2.19, 1.13]
2.33 Expressive language, adjusted mean difference (months) in improvement in Schlichting test score (equivalent age - real age): medium-term	1		Mean Difference (IV, Random, 95% CI)	0.96 [-0.43, 2.35]
2.34 Non-word repetition total score, adjusted OR (long-term)	1		Odds Ratio (IV, Random, 95% CI)	1.69 [0.64, 4.47]
2.35 Reading , WORD test, adjusted OR (long-term)	1		Odds Ratio (IV, Random, 95% CI)	Subtotals only
2.36 Spelling, ALSPAC test, adjusted OR (long-term)	1		Odds Ratio (IV, Random, 95% CI)	0.90 [0.33, 2.45]
2.37 Phoneme deletion, adjusted OR (long-term)	1		Odds Ratio (IV, Random, 95% CI)	0.84 [0.32, 2.20]
2.38 Cognitive development: Griffiths practical reasoning (medium-term)	1		Mean Difference (IV, Random, 95% CI)	2.40 [-3.78, 8.58]
2.39 Cognitive development: IQ (WISC-III UK short form) adjusted OR (long term)	1		Odds Ratio (IV, Random, 95% CI)	Totals not selected
2.40 Behaviour, Richman score (medium-term)	1	150	Mean Difference (IV, Random, 95% CI)	-0.65 [-1.85, 0.55]
2.41 Behaviour, Richman score, dichotomised (medium-term)	1	150	Risk Ratio (M-H, Random, 95% CI)	0.63 [0.42, 0.96]
2.42 Behaviour, Richman score, adjusted OR (medium-term)	1		Odds Ratio (IV, Random, 95% CI)	1.16 [0.27, 4.90]
2.43 Behaviour, Richman score (long-term)	1	123	Mean Difference (IV,	0.90 [-0.27, 2.07]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
			Random, 95% CI)	
2.44 Behaviour, Richman score, dichotomised (long-term)	1	123	Risk Ratio (M-H, Random, 95% CI)	1.22 [0.62, 2.40]
2.45 Behaviour: SDQ teacher report, total, adjusted OR (long-term)	1		Odds Ratio (IV, Random, 95% CI)	2.05 [0.62, 6.74]
2.46 Parent-child interaction: Erickson child scale (medium-term)	1	165	Mean Difference (IV, Random, 95% CI)	-0.34 [-0.56, -0.12]
2.47 Parent-child interaction: Erickson parent scale (medium-term)	1	165	Mean Difference (IV, Random, 95% CI)	-0.42 [-0.67, -0.17]
2.48 Parental stress, Parental Stress Index, short form (long-term)	1	383	Mean Difference (IV, Random, 95% CI)	0.00 [-4.12, 4.12]
2.49 Generic health-related quality of life: TAIQOL (medium-term)	1		Mean Difference (IV, Fixed, 95% CI)	Subtotals only
2.49.1 Vitality	1	165	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-1.95, 1.75]
2.49.2 Appetite	1	165	Mean Difference (IV, Fixed, 95% CI)	0.40 [-3.77, 4.57]
2.49.3 Communication	1	165	Mean Difference (IV, Fixed, 95% CI)	0.30 [-5.11, 5.71]
2.49.4 Motoric	1	165	Mean Difference (IV, Fixed, 95% CI)	0.00 [-2.51, 2.51]
2.49.5 Social	1	165	Mean Difference (IV, Fixed, 95% CI)	0.00 [-2.49, 2.49]
2.49.6 Anxiety	1	165	Mean Difference (IV, Fixed, 95% CI)	0.30 [-3.04, 3.64]
2.49.7 Aggression	1	165	Mean Difference (IV, Fixed, 95% CI)	0.30 [-5.82, 6.42]
2.49.8 Eating	1	165	Mean Difference (IV, Fixed, 95% CI)	-0.10 [-1.63, 1.43]
2.49.9 Sleeping	1	165	Mean Difference (IV, Fixed, 95% CI)	0.00 [-5.70, 5.70]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.50 Literacy (long-term)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.50.1 Woodcock Reading Mastery Tests: Word identification subtest	1	391	Mean Difference (IV, Random, 95% CI)	-1.00 [-3.28, 1.28]
2.50.2 Woodcock Reading Mastery Tests: Word Attack subtest	1	391	Mean Difference (IV, Random, 95% CI)	-1.00 [-3.68, 1.68]
2.50.3 Woodcock Reading Mastery Tests: Passage Comprehension subtest	1	391	Mean Difference (IV, Random, 95% CI)	-1.00 [-3.38, 1.38]
2.50.4 Oral reading fluency test: Children in grade 3	1	74	Mean Difference (IV, Random, 95% CI)	-9.00 [-26.58, 8.58]
2.50.5 Oral reading fluency test: Children in grade 4	1	184	Mean Difference (IV, Random, 95% CI)	0.00 [-10.70, 10.70]
2.50.6 Oral reading fluency test: Children in grade 5	1	105	Mean Difference (IV, Random, 95% CI)	-5.00 [-18.98, 8.98]
2.50.7 Oral reading fluency test: Children in grade 6	1	21	Mean Difference (IV, Random, 95% CI)	6.00 [-27.42, 39.42]
2.50.8 Woodcock–Johnson III Tests of Achievement: Spelling subtest	1	390	Mean Difference (IV, Random, 95% CI)	-1.00 [-3.89, 1.89]
2.50.9 Woodcock–Johnson III Tests of Achievement: Writing Samples subtest	1	387	Mean Difference (IV, Random, 95% CI)	-1.00 [-3.89, 1.89]
2.51 Phonological awareness (long-term)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.51.1 Comprehensive Test of Phonological Processing: Elision subtest	1	391	Mean Difference (IV, Random, 95% CI)	-0.10 [-0.91, 0.71]
2.51.2 Comprehensive Test of Phonological Processing: Rapid Letter Naming subtest	1	389	Mean Difference (IV, Random, 95% CI)	-0.30 [-0.79, 0.19]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.52 Attention, impulsivity, and psychosocial function, long-term (1): disruptive behaviour disorders, child behaviour and impairment rating	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.52.1 Disruptive Behavior Disorders Rating Scale: Inattention factor: Parent's rating	1	390	Mean Difference (IV, Random, 95% CI)	0.05 [-0.08, 0.18]
2.52.2 Disruptive Behavior Disorders Rating Scale: Inattention factor: Teacher's rating	1	382	Mean Difference (IV, Random, 95% CI)	0.04 [-0.11, 0.19]
2.52.3 Disruptive Behavior Disorders Rating Scale: Impulsivity and overactivity factor: Parent's rating	1	390	Mean Difference (IV, Random, 95% CI)	0.10 [-0.01, 0.21]
2.52.4 Disruptive Behavior Disorders Rating Scale: Impulsivity and overactivity factor: Teacher's rating	1	382	Mean Difference (IV, Random, 95% CI)	0.08 [-0.04, 0.20]
2.52.5 Disruptive Behavior Disorders Rating Scale: Oppositional defiant factor: Parent's rating	1	390	Mean Difference (IV, Random, 95% CI)	0.05 [-0.06, 0.16]
2.52.6 Disruptive Behavior Disorders Rating Scale: Oppositional defiant factor: Teacher's rating	1	382	Mean Difference (IV, Random, 95% CI)	0.00 [-0.11, 0.11]
2.52.7 Child Behavior Checklist: Total Problems score, parent’s rating	1	390	Mean Difference (IV, Random, 95% CI)	2.00 [-0.38, 4.38]
2.52.8 Child Behavior Checklist: Total Problems score, teacher’s rating	1	380	Mean Difference (IV, Random, 95% CI)	2.00 [-0.21, 4.21]
2.52.9 Impairment Rating Scales: Overall functioning, parent’s rating	1	390	Mean Difference (IV, Random, 95% CI)	0.14 [-0.13, 0.41]
2.52.10 Impairment Rating Scales: Overall functioning, teacher’s rating	1	382	Mean Difference (IV, Random, 95% CI)	0.26 [-0.18, 0.70]
2.53 Attention, impulsivity, and psychosocial function, long-term (2): social skills	1		Mean Difference (IV, Random, 95% CI)	Subtotals only

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
2.53.1 Attention, impulsivity, and psychosocial function: Social Skills Rating System: parent version	1	388	Mean Difference (IV, Random, 95% CI)	-2.00 [-5.68, 1.68]
2.53.2 Attention, impulsivity, and psychosocial function: Social Skills Rating System: teacher version	1	370	Mean Difference (IV, Random, 95% CI)	-1.00 [-3.65, 1.65]
2.54 Attention, impulsivity, and psychosocial function, long-term: Visual and auditory continuous performance	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.54.1 Visual Continuous Performance Test: Inattention	1	391	Mean Difference (IV, Random, 95% CI)	0.20 [-2.66, 3.06]
2.54.2 Visual Continuous Performance Test: Impulsivity	1	391	Mean Difference (IV, Random, 95% CI)	0.60 [-2.58, 3.78]
2.54.3 Auditory Continuous Performance Test: Inattention	1	308	Mean Difference (IV, Random, 95% CI)	-0.30 [-2.00, 1.40]
2.54.4 Auditory Continuous Performance Test: Impulsivity	1	307	Mean Difference (IV, Random, 95% CI)	-0.90 [-3.26, 1.46]
2.55 Intelligence and academic achievement (long-term)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only
2.55.1 Wechsler Abbreviated Scale of Intelligence	1	391	Mean Difference (IV, Random, 95% CI)	0.00 [-2.68, 2.68]
2.55.2 Calculation subtest of the Woodcock–Johnson III Tests of Achievement	1	389	Mean Difference (IV, Random, 95% CI)	0.00 [-2.58, 2.58]

Comparison 3

VT versus non-surgical treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
3.1 Mean final hearing threshold (short-term)	1	125	Mean Difference (IV, Random, 95% CI)	-9.00 [-12.61, -5.39]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size	
3.2 Mean final hearing threshold (medium-term)	1	125	Mean Difference (IV, Random, 95% CI)	-5.98 [-9.21, -2.75]	
3.3 Adverse event: myringosclerosis (long-term)	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected	
3.4 Number of doctor-diagnosed AOM episodes (medium-term)	1	125	Mean Difference (IV, Random, 95% CI)	-0.23 [-0.42, -0.04]	
3.5 Number of doctor-diagnosed episodes of AOM (long-term)	1	125	Mean Difference (IV, Random, 95% CI)	-0.05 [-0.31, 0.21]	

Comparison 4

VT versus myringotomy

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size	
4.1 Hearing returned to normal: VT versus laser myringotomy (medium-term)	2		Risk Ratio (M-H, Random, 95% CI)	Subtotals only	
4.1.1 Adjusted for non-independence of within-individual measurements. Assumed ICC of 0.5	2	132	Risk Ratio (M-H, Random, 95% CI)	1.22 [0.59, 2.53]	
4.2 Mean final hearing threshold, randomised by child (short-term). Adjusted for non-independence of within-individual measurements. Assumed ICC of 0.5	1	104	Mean Difference (IV, Random, 95% CI)	0.20 [-2.13, 2.53]	
4.3 Mean final hearing threshold, randomised by ear (short-term)	1		Mean Difference (IV, Random, 95% CI)	Subtotals only	
4.4 Mean final hearing threshold (medium-term)	2		Mean Difference (IV, Fixed, 95% CI)	Subtotals only	
4.4.1 Pure tone audiometry at 12 months.	1	104	Mean Difference	0.80 [-0.87, 2.47]	

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
Adjusted for non-independence of within-individual measurements, assumed ICC 0.5			(IV, Fixed, 95% CI)	
4.4.2 Air bone gap at 12 months.	1	50	Mean Difference (IV, Fixed, 95% CI)	4.50 [0.76, 8.24]
4.5 Adverse event: persistent perforation (medium-term)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
4.5.1 Adjustment for non-independence of within-individual measurements: Assumed ICC of 0.5	1	102	Risk Ratio (M-H, Random, 95% CI)	1.00 [0.06, 15.56]
4.6 Adverse event: persistent perforation cold-steel myringotomy (medium-term)	2	208	Peto Odds Ratio (Peto, Fixed, 95% CI)	8.09 [1.78, 36.79]
4.7 Persistence of OME: VT versus laser myringotomy (short-term)	1	102	Risk Ratio (M-H, Random, 95% CI)	1.40 [0.48, 4.12]
4.7.1 Adjusted for non-independence of within-individual measurements. Assumed ICC of 0.5	1	102	Risk Ratio (M-H, Random, 95% CI)	1.40 [0.48, 4.12]
4.8 Persistence of OME: VT versus thermal myringotomy, randomised by ear (short-term)	1	72	Peto Odds Ratio (Peto, Fixed, 95% CI)	0.11 [0.02, 0.53]
4.9 Persistence of OME: VT versus cold-steel myringotomy (medium-term)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
4.10 Persistence of OME: VT versus laser myringotomy (medium-term)	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
4.10.1 Adjusted for non-	1		Risk Ratio (M-H,	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
independence of within-participant measurements: Assumed ICC of 0.5			Random, 95% CI)	
4.11 Persistence of OME: VT versus laser myringotomy, randomised by ear (medium-term)	1		Odds Ratio (IV, Random, 95% CI)	Totals not selected
4.11.1 Correlation coefficient of 0.5 assumed	1		Odds Ratio (IV, Random, 95% CI)	Totals not selected
4.12 Persistence of OME: mean days to first recurrence	1	389	Mean Difference (IV, Random, 95% CI)	173.88 [150.19, 197.56]
4.13 Persistence of OME (long-term)	1	491	Risk Ratio (M-H, Random, 95% CI)	0.97 [0.90, 1.05]
4.14 Adverse events: otorrhoea (long-term)	1	491	Risk Ratio (M-H, Random, 95% CI)	1.58 [0.98, 2.53]
4.15 Zero, one or two episodes of AOM in 12 months	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
4.15.1 Zero episodes	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
4.15.2 One episode	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
4.15.3 Two episodes	1		Risk Ratio (M-H, Random, 95% CI)	Totals not selected
4.16 Three or more episodes of AOM in 12 months	1		Peto Odds Ratio (Peto, Fixed, 95% CI)	Totals not selected
4.16.1 Three episodes	1		Peto Odds Ratio (Peto, Fixed, 95% CI)	Totals not selected
4.16.2 Four or more episodes	1		Peto Odds	Totals not selected

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
			Ratio (Peto, Fixed, 95% CI)	
4.17 Adverse event: retraction of TM: VT versus laser myringotomy (medium-term)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
4.17.1 Adjusted for non-independence of within-individual measurements. Assumed ICC of 0.5	1	102	Risk Ratio (M-H, Random, 95% CI)	2.67 [0.75, 9.48]
4.18 Adverse event: hypertrophic scar of TM: VT versus laser myringotomy (medium-term)	1		Peto Odds Ratio (Peto, Fixed, 95% CI)	Subtotals only
4.19 Adverse event: otorrhoea: VT versus laser myringotomy (medium-term)	1		Risk Ratio (M-H, Random, 95% CI)	Subtotals only
4.19.1 Adjusted for non-independence of within-individual measurements: assumed ICC of 0.5	1	102	Risk Ratio (M-H, Random, 95% CI)	4.00 [0.46, 34.57]

Comparison 5

Sensitivity analyses: VT versus no treatment

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
5.1 Sensitivity analysis: Return to normal hearing, randomised by ear (medium-term). CC 0.3	1		Odds Ratio (IV, Random, 95% CI)	1.13 [0.46, 2.74]
5.1.1 Sensitivity analysis: normal defined as <15dB. CC=0.3	1		Odds Ratio (IV, Random, 95% CI)	1.13 [0.46, 2.74]
5.2 Sensitivity analysis. Return to normal hearing, randomised by ear (medium-term). CC 0.7	1		Odds Ratio (IV, Random, 95% CI)	1.13 [0.47, 2.75]
5.2.1 Sensitivity analysis: normal	1		Odds Ratio (IV,	1.13 [0.47, 2.75]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
defined as <15dB. CC=0.7			Random, 95% CI)	
5.3 Sensitivity analysis. Return to normal hearing, randomised by ear (medium-term). Normal defined as <25dB. CC=0.5	1		Odds Ratio (IV, Random, 95% CI)	1.00 [0.57, 1.76]
5.3.1 Sensitivity analysis: normal defined as <25dB. CC=0.5 (medium-term)	1		Odds Ratio (IV, Random, 95% CI)	1.00 [0.57, 1.76]
5.4 Sensitivity analysis. Mean final hearing threshold, randomised by ear (medium-term). CC0.3	2		Mean Difference (IV, Random, 95% CI)	-3.47 [-10.01, 3.06]
5.4.1 Sensitivity analysis: correlation coefficient = 0.3	2		Mean Difference (IV, Random, 95% CI)	-3.47 [-10.01, 3.06]
5.5 Sensitivity analysis. Mean final hearing threshold, randomised by ear (medium-term). CC0.7	2		Mean Difference (IV, Random, 95% CI)	-3.49 [-10.37, 3.38]
5.5.1 Sensitivity analysis: correlation coefficient = 0.7	2		Mean Difference (IV, Random, 95% CI)	-3.49 [-10.37, 3.38]
5.6 Sensitivity analysis. Change in hearing threshold from baseline, randomised by ear (medium-term). CC0.3	1		Mean Difference (IV, Random, 95% CI)	-0.10 [-3.22, 3.01]
5.7 Sensitivity analysis. Change in hearing threshold from baseline, randomised by ear (medium-term). CC0.7	1		Mean Difference (IV, Random, 95% CI)	-0.21 [-3.34, 2.92]
5.8 Sensitivity analysis. Adverse event: perforation/retraction, randomised by ear (medium-term). CC=0.3	1		Odds Ratio (IV, Random, 95% CI)	0.85 [0.33, 2.21]
5.8.1 Sensitivity analysis: correlation coefficient 0.3	1		Odds Ratio (IV, Random, 95% CI)	0.85 [0.33, 2.21]
5.9 Sensitivity analysis. Adverse event: perforation/retraction, randomised by ear	1		Odds Ratio (IV, Random, 95% CI)	0.91 [0.45, 1.86]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size	
(medium-term). CC=0.7					
5.9.1 Sensitivity analysis: correlation coefficient 0.7	1		Odds Ratio (IV, Random, 95% CI)	0.91 [0.45, 1.86]	
5.10 Sensitivity analysis. Persistence of OME: randomised by child (medium-term). ICC 1.0	1	40	Risk Ratio (M-H, Random, 95% CI)	0.27 [0.11, 0.70]	
5.10.1 Sensitivity analysis: assuming ICC of 1.0 (complete correlation between ears)	1	40	Risk Ratio (M-H, Random, 95% CI)	0.27 [0.11, 0.70]	
5.11 Sensitivity analysis. Persistence of OME: randomised by child (medium-term). ICC zero	1	81	Risk Ratio (M-H, Random, 95% CI)	0.30 [0.16, 0.56]	
5.11.1 Sensitivity analysis: assuming ICC of 0.0 (no correlation between ears)	1	81	Risk Ratio (M-H, Random, 95% CI)	0.30 [0.16, 0.56]	
5.12 Sensitivity analysis. Persistence of OME: randomised by ear (medium-term). CC 0.3	1		Odds Ratio (IV, Random, 95% CI)	0.66 [0.24, 1.83]	
5.12.1 Sensitivity analysis: correlation coefficient = 0.3	1		Odds Ratio (IV, Random, 95% CI)	0.66 [0.24, 1.83]	
5.13 Sensitivity analysis. Persistence of OME: randomised by ear (medium-term). CC 0.7	1		Odds Ratio (IV, Random, 95% CI)	0.66 [0.24, 1.83]	
5.13.1 Sensitivity analysis: correlation coefficient = 0.7	1		Odds Ratio (IV, Random, 95% CI)	0.66 [0.24, 1.83]	

Comparison 6

Sensitivity analyses: Early VT versus watchful waiting

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.1 Sensitivity analysis. Mean final hearing threshold (air-bone gap), randomised by child, analysed by ear (medium-term). ICC 1.0	1	87	Mean Difference (IV, Random, 95% CI)	-1.18 [-3.08, 0.72]
6.1.1 Sensitivity analysis: assuming ICC of 1.0 (complete correlation between ears)	1	87	Mean Difference (IV, Random, 95% CI)	-1.18 [-3.08, 0.72]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.2 Sensitivity analysis. Mean final hearing threshold (air-bone gap), randomised by child, analysed by ear (medium-term). ICC zero	1	160	Mean Difference (IV, Random, 95% CI)	-1.18 [-2.58, 0.22]
6.2.1 Sensitivity analysis: assuming ICC of 0.0 (no correlation between ears)	1	160	Mean Difference (IV, Random, 95% CI)	-1.18 [-2.58, 0.22]
6.3 Sensitivity analysis. Mean final hearing threshold, randomised by child (long-term). CC for Paradise 2007 of 0.3	3	633	Mean Difference (IV, Random, 95% CI)	0.37 [-0.37, 1.11]
6.3.1 Sensitivity analysis: cc for Paradise 2007 (left and right ear data combined) of 0.3	3	633	Mean Difference (IV, Random, 95% CI)	0.37 [-0.37, 1.11]
6.4 Sensitivity analysis. Mean final hearing threshold, randomised by child (long-term). CC for Paradise 2007 of 0.7	3	633	Mean Difference (IV, Random, 95% CI)	0.35 [-0.45, 1.16]
6.4.1 Sensitivity analysis: cc for Paradise 2007 (left and right ear data combined) of 0.7	3	633	Mean Difference (IV, Random, 95% CI)	0.35 [-0.45, 1.16]
6.5 Sensitivity analysis. Persistent perforation, randomised by child (long-term). ICC 1.0	1	281	Risk Ratio (M-H, Random, 95% CI)	2.73 [0.29, 25.97]
6.5.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)	1	281	Risk Ratio (M-H, Random, 95% CI)	2.73 [0.29, 25.97]
6.6 Sensitivity analysis. Persistent perforation, randomised by child (long-term). ICC zero	1	562	Risk Ratio (M-H, Fixed, 95% CI)	2.73 [0.56, 13.43]
6.6.1 Sensitivity analysis: ICC zero (no correlation between ears)	1	562	Risk Ratio (M-H, Fixed, 95% CI)	2.73 [0.56, 13.43]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
6.7 Sensitivity analysis. Persistence of OME, randomised by child, measured by otoscopy (medium-term). ICC 1.0	1	87	Risk Ratio (M-H, Random, 95% CI)	0.49 [0.11, 2.22]
6.7.1 Sensitivity analysis: assuming ICC of 1.0 (complete correlation between ears)	1	87	Risk Ratio (M-H, Random, 95% CI)	0.49 [0.11, 2.22]
6.8 Sensitivity analysis. Persistence of OME, randomised by child, measured by otoscopy (medium-term). ICC=zero	1	161	Risk Ratio (M-H, Random, 95% CI)	0.40 [0.12, 1.34]
6.8.1 Sensitivity analysis: assuming ICC of 0.0 (no correlation between ears)	1	161	Risk Ratio (M-H, Random, 95% CI)	0.40 [0.12, 1.34]
6.9 Sensitivity analysis. Tympanosclerosis (long term). ICC=1.0	1	281	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.27, 3.08]
6.9.1 Sensitivity analysis: ICC 1.0 (full correlation between ears)	1	281	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.27, 3.08]
6.10 Sensitivity analysis. Tympanosclerosis (long term). ICC=zero	1	562	Risk Ratio (M-H, Random, 95% CI)	0.83 [0.36, 1.92]
6.10.1 Sensitivity analysis ICC zero (no correlation between ears)	1	562	Risk Ratio (M-H, Random, 95% CI)	0.83 [0.36, 1.92]
6.11 Sensitivity analysis. Adverse event: fibrosis (long term). ICC=1.0	1	281	Risk Ratio (M-H, Random, 95% CI)	0.46 [0.04, 4.97]
6.11.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)	1	281	Risk Ratio (M-H, Random, 95% CI)	0.46 [0.04, 4.97]
6.12 Sensitivity analysis. Adverse event: fibrosis (long term). ICC=zero	1	562	Risk Ratio (M-H, Random, 95% CI)	0.68 [0.15, 3.03]
6.12.1 Sensitivity analysis: ICC	1	562	Risk Ratio (M-	0.68 [0.15, 3.03]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
zero (no correlation between ears)			H, Random, 95% CI)	
6.13 Sensitivity analysis. Segmental atrophy (long term). ICC=1.0	1	281	Risk Ratio (M-H, Random, 95% CI)	2.92 [1.72, 4.96]
6.13.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)	1	281	Risk Ratio (M-H, Random, 95% CI)	2.92 [1.72, 4.96]
6.14 Sensitivity analysis. Segmental atrophy (long term). ICC=zero	1	562	Risk Ratio (M-H, Random, 95% CI)	2.85 [1.97, 4.13]
6.14.1 Sensitivity analysis: ICC zero (no correlation between ears)	1	562	Risk Ratio (M-H, Random, 95% CI)	2.85 [1.97, 4.13]
6.15 Sensitivity analysis. Retraction pocket with other abnormality (long term). ICC=1.0	1	281	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.06, 14.43]
6.15.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)	1	281	Risk Ratio (M-H, Random, 95% CI)	0.91 [0.06, 14.43]
6.16 Sensitivity analysis. Retraction pocket with other abnormality (long term). ICC=zero	1	562	Odds Ratio (M-H, Fixed, 95% CI)	0.91 [0.06, 14.64]
6.16.1 Sensitivity analysis: ICC zero (no correlation between ears)	1	562	Odds Ratio (M-H, Fixed, 95% CI)	0.91 [0.06, 14.64]
6.17 Sensitivity analysis. Parent-child interaction: Erickson child scale (medium-term). CC0.3	1	165	Mean Difference (IV, Random, 95% CI)	-0.34 [-0.53, -0.15]
6.18 Sensitivity analysis. Parent-child interaction: Erickson child scale (medium-term). CC0.7	1	165	Mean Difference (IV, Random, 95% CI)	-0.34 [-0.58, -0.10]
6.19 Sensitivity analysis. Parent-child interaction: Erickson parent scale (medium-term). CC0.3	1	165	Mean Difference (IV, Random, 95% CI)	-0.42 [-0.64, -0.20]
6.20 Sensitivity analysis. Parent-child interaction: Erickson parent	1	165	Mean Difference (IV,	-0.42 [-0.70, -0.14]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size	
scale (medium-term). CC=0.7			Random, 95% CI)		

Comparison 7

Sensitivity analyses: VT versus myringotomy

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.1 Sensitivity analysis. Hearing returned to normal: VT versus laser myringotomy (medium-term). ICC=1.0	2	112	Risk Ratio (M-H, Random, 95% CI)	1.21 [0.59, 2.48]
7.1.1 Sensitivity analysis: ICC of 1.0 (complete correlation between ears)	2	112	Risk Ratio (M-H, Random, 95% CI)	1.21 [0.59, 2.48]
7.2 Sensitivity analysis. Hearing returned to normal: VT versus laser myringotomy (medium-term). ICC=zero	2	166	Risk Ratio (M-H, Random, 95% CI)	1.22 [0.62, 2.40]
7.2.1 Sensitivity analysis: ICC of zero (no correlation between ears)	2	166	Risk Ratio (M-H, Random, 95% CI)	1.22 [0.62, 2.40]
7.3 Sensitivity analysis. Mean final hearing threshold, randomised by child (short-term). ICC 1.0	1	78	Mean Difference (IV, Random, 95% CI)	0.20 [-2.50, 2.90]
7.4 Sensitivity analysis. Mean final hearing threshold, randomised by child (short-term). ICC=zero	1	156	Mean Difference (IV, Random, 95% CI)	0.20 [-1.71, 2.11]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
7.5 Sensitivity analysis. Mean final hearing threshold (medium-term). ICC=1.0	1	78	Mean Difference (IV, Random, 95% CI)	0.80 [-1.13, 2.73]
7.5.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)	1	78	Mean Difference (IV, Random, 95% CI)	0.80 [-1.13, 2.73]
7.6 Sensitivity analysis. Mean final hearing threshold (medium-term). ICC=zero	1	156	Mean Difference (IV, Random, 95% CI)	0.80 [-0.57, 2.17]
7.6.1 Sensitivity analysis: ICC zero (no correlation between ears)	1	156	Mean Difference (IV, Random, 95% CI)	0.80 [-0.57, 2.17]
7.7 Sensitivity analysis. Persistent perforation (medium-term). ICC=1.0	1	82	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.06, 15.45]
7.7.1 Sensitivity analysis: ICC = 1 (complete correlation between ears)	1	82	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.06, 15.45]
7.8 Sensitivity analysis. Persistent perforation (medium-term). ICC=zero	1	136	Risk Ratio (M-H, Random, 95% CI)	2.00 [0.19, 21.54]
7.8.1 Sensitivity analysis: ICC of zero (no correlation between ears)	1	136	Risk Ratio (M-H, Random, 95% CI)	2.00 [0.19, 21.54]
7.9 Sensitivity analysis.	1	82	Risk Ratio (M-H,	1.50 [0.46, 4.92]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
Persistence of OME: VT versus laser myringotomy (short-term). ICC=1.0			Random, 95% CI)	
7.9.1 Sensitivity analysis: ICC of 1.0 (complete correlation between ears)	1	82	Risk Ratio (M-H, Random, 95% CI)	1.50 [0.46, 4.92]
7.10 Sensitivity analysis. Persistence of OME: VT versus laser myringotomy (short-term) ICC=zero	1	136	Risk Ratio (M-H, Random, 95% CI)	1.43 [0.58, 3.53]
7.10.1 Sensitivity analysis: ICC of zero (no correlation between ears)	1	136	Risk Ratio (M-H, Random, 95% CI)	1.43 [0.58, 3.53]
7.11 Sensitivity analysis. Persistence of OME: VT versus laser myringotomy (medium-term). ICC=1.0	1	82	Risk Ratio (M-H, Random, 95% CI)	0.35 [0.17, 0.74]
7.11.1 Sensitivity analysis: ICC of 1.0 (complete correlation between ears)	1	82	Risk Ratio (M-H, Random, 95% CI)	0.35 [0.17, 0.74]
7.12 Sensitivity analysis. Persistence of OME: VT versus laser myringotomy (medium-term). ICC=zero	1	136	Risk Ratio (M-H, Random, 95% CI)	0.33 [0.18, 0.60]
7.12.1 Sensitivity analysis: ICC of zero (no correlation between ears)	1	136	Risk Ratio (M-H, Random, 95% CI)	0.33 [0.18, 0.60]
7.13 Sensitivity	1		Odds Ratio (IV,	0.27 [0.18, 0.42]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size
analysis. Persistence of OME: VT versus laser myringotomy, randomised by ear (medium-term). CC=0.3			Random, 95% CI)	
7.13.1 Sensitivity analysis: correlation coefficient of 0.3 assumed	1		Odds Ratio (IV, Random, 95% CI)	0.27 [0.18, 0.42]
7.14 Sensitivity analysis. Persistence of OME: VT versus laser myringotomy, randomised by ear (medium-term). CC=0.7	1		Odds Ratio (IV, Random, 95% CI)	0.27 [0.21, 0.36]
7.14.1 Sensitivity analysis: correlation coefficient of 0.7 assumed	1		Odds Ratio (IV, Random, 95% CI)	0.27 [0.21, 0.36]
7.15 Sensitivity analysis. Retraction of TM: VT versus laser myringotomy (medium-term). ICC=1.0	1	82	Risk Ratio (M-H, Random, 95% CI)	3.50 [0.77, 15.85]
7.15.1 Sensitivity analysis: ICC of 1.0 (complete correlation between ears)	1	82	Risk Ratio (M-H, Random, 95% CI)	3.50 [0.77, 15.85]
7.16 Sensitivity analysis. Retraction of TM: VT versus laser myringotomy (medium-term). ICC=zero	1	136	Risk Ratio (M-H, Random, 95% CI)	2.75 [0.92, 8.21]
7.16.1 Sensitivity analysis: ICC of zero (no correlation	1	136	Risk Ratio (M-H, Random, 95% CI)	2.75 [0.92, 8.21]

Outcome or subgroup title	No. of studies	No. of participants	Statistical method	Effect size	
between ears)					
7.17 Sensitivity analysis. Otorrhoea: VT versus laser myringotomy (medium-term). ICC=1.0	1	82	Risk Ratio (M-H, Random, 95% CI)	3.00 [0.33, 27.66]	
7.17.1 Sensitivity analysis: 1.0 (complete correlation between ears)	1	82	Risk Ratio (M-H, Random, 95% CI)	3.00 [0.33, 27.66]	
7.18 Sensitivity analysis. Otorrhoea: VT versus laser myringotomy (medium-term). ICC=zero	1	136	Risk Ratio (M-H, Random, 95% CI)	2.50 [0.50, 12.44]	
7.18.1 Sensitivity analysis: ICC of zero (no correlation between ears)	1	136	Risk Ratio (M-H, Random, 95% CI)	2.50 [0.50, 12.44]	

History

Protocol first published: Issue 3, 2022

Contributions of authors

Samuel MacKeith: drafted the protocol. Screened the search results and selected studies. Reviewed the analyses and reviewed and edited the text of the review.

Caroline A Mulvaney: drafted the protocol. Screened the search results and selected studies, conducted data extraction, carried out statistical analyses and GRADE assessment. Drafted the text of the review.

Kevin Galbraith: drafted the protocol. Screened the search results and selected studies, conducted data extraction, carried out statistical analyses and GRADE assessment. Drafted the text of the review.

Katie Webster: Screened the search results and selected studies. Drafted the text of the review.

Rachel Connolly: Conducted data extraction. Reviewed the analyses and reviewed and edited the text of the review.

Aye Paing: Conducted data extraction. Reviewed the analyses and reviewed and edited the text of the review.

Tal Marom: reviewed the protocol. Reviewed the analyses and reviewed and edited the text of the review.

Mat Daniel: reviewed the protocol. Reviewed the analyses and reviewed and edited the text of the review.

Roderick P Venekamp: co-wrote and edited the protocol. Reviewed the analyses and reviewed and edited the text of the review.

Maroeska Rovers: co-wrote and edited the protocol. Reviewed the analyses and reviewed and edited the text of the review.

Anne GM Schilder: co-wrote and edited the protocol. Reviewed the analyses and reviewed and edited the text of the review.

Declarations of interest

Samuel MacKeith: treats patients with OME in his NHS and private practice and is Assistant Co-ordinating Editor of Cochrane ENT but has not been involved in the editorial process for this review.

Caroline A Mulvaney: none known.

Kevin Galbraith: none known.

Katie Webster: none known.

Rachel Connolly:

Aye Paing:

Tal Marom: has no conflict of interests to declare.

Mat Daniel: has a financial interest in Aventamed, a company that produces a ventilation tube insertion device.

Roderick P Venekamp: is an Editor for the Cochrane Acute Respiratory Infections Group and Cochrane ENT, but had no role in the editorial process for this review.

Maroeska M Rovers: has no financial conflicts of interest. She has previously performed a randomised controlled trial and individual patient data meta-analysis on the effect of ventilation tubes, and has acted as a member of the Dutch guideline committee on otitis media.

Anne GM Schilder: Professor Anne Schilder was joint Co-ordinating Editor of Cochrane ENT until April 2020, but had no role in the editorial process for this review. Her evidENT team at the UCL Ear Institute is supported by the National Institute of Health Research (NIHR) University College London Hospitals (UCLH) Biomedical Research Centre (BRC), with research projects being supported by the NIHR, Wellcome Trust, RNiD, ENT UK and industry. She is the National Specialty Lead for the NIHR Clinical Research Network ENT and Surgical Specialty Lead for ENT for the Royal College of Surgeons of England's Clinical Research Initiative. In her role as director of the NIHR UCLH BRC Deafness and Hearing Problems Theme, she advises CRO, biotech and pharma companies in the hearing field on clinical trial design and delivery.

Sources of support

Internal sources

- No sources of support provided

External sources

- National Institute for Health Research, UK
  Infrastructure funding for Cochrane ENT

Differences between protocol and review

In the protocol for this review we planned to assess the following six comparisons (MacKeith 2022b):

- bilateral ventilation tubes versus no treatment/watchful waiting;
- bilateral ventilation tubes versus hearing aids;
- bilateral ventilation tubes versus non-surgical treatment;
- bilateral ventilation tubes versus myringotomy alone;
- unilateral ventilation tubes versus no treatment/watchful waiting;
- unilateral ventilation tubes versus myringotomy alone in the other ear/other children.

However, two issues arose whilst conducting the review. Firstly, we agreed that the comparators 'no treatment' and 'watchful waiting' for this review were different. No treatment indicates that it was intended that children in the comparator arm would not receive treatment during the study. Watchful waiting suggests a more active follow-up, with intervention at a later stage as required. We therefore considered it appropriate to separate these comparisons.

The second issue was that studies often included a mixture of children with unilateral and bilateral OME, therefore the distinction between unilateral and bilateral ventilation tube insertion was not relevant.

We therefore revised our comparisons to the following:

- ventilation tubes (bilateral or unilateral) versus no treatment
- early ventilation tubes versus watchful waiting (treatment later if required);
- ventilation tubes versus hearing aids;
- ventilation tubes versus non-surgical treatment;
- ventilation tubes versus myringotomy alone.

Characteristics of studies

Characteristics of included studies [ordered by study ID]

Bernard 1991

Study characteristics	
Methods	Single centre, parallel group RCT with 18 months of follow-up. Randomised by child.
Participants	**Location:** Canada, single centre **Setting of recruitment and treatment:** Otolaryngology clinic at the Children's Hospital of Eastern Ontario. **Study dates:** Not reported **Sample size:** • **Number randomised:** 139 (68 to surgical treatment; 71 to medical treatment) • **Number completed:** 125 (60 in surgical treatment group; 65 in medical treatment group) **Participant (baseline) characteristics:** Age, years: • Surgical treatment = mean 4.7 years • Medical treatment = mean 5.0 years Gender

	• Surgical treatment: 34 (56.7%) male: 26 (43.3%) female • Medical treatment: 34 (52.3%) male: 31 (47.7%) female Hearing loss at baseline • Surgical treatment = mean 30.7 db HL • Medical treatment = mean 29.6 db HL **Inclusion criteria:** • Age 2.5 to 7 years; • Long-standing (greater than 3 months) middle ear effusion as indicated by type "B" tympanogram (in at least one ear) and otoscopic evidence (fluid/air fluid levels) of middle ear effusion for at least 3 months preceding entry into the trial; • At least two physician-documented trials of antibacterials for AOM or OME, of at least 10 days' duration in the 3 months preceding entry into the trial; • History of hearing loss (based on parental reports) of >3 months' duration; at the time of entry into the trial: • Hearing loss of at least 25 dB HL (hearing level based on the ANSI 53.6 1969 standard) air conduction at 2 or more frequencies 0.5, 1, 2, and 4kHz (pure-tone audiometry) in at least one ear; • Bone conduction thresholds within normal limits (0 to 10 db HL) bilaterally; • Air-bone gap of >15dB at frequencies with elevated air conduction thresholds. **Exclusion criteria:** • cervicofacial abnormality (cleft palate, Down syndrome); • documented immune insufficiency; • documented allergy to sulfonamide; • previous insertion of VT; • documented speech delay.
Interventions	**Intervention** Bilateral myringotomy and insertion of VTs at the anterior-inferior quadrant of the tympanic membrane by the same otolaryngologist. n=68 **Comparison** Sulfisoxazole, 75mg/kg divided into 2 daily doses for 6 months n=71
Outcomes	Proportion with normal/impaired hearing (not extracted because of insufficient data.) Mean final hearing threshold • Assessed with pure tone audiometry at 0.5, 1, 2 and 4kHz Adverse events: • Persistent perforation • Myringosclerosis • Tube otorrhoea • Antibiotic group: medication related side effects, rash, nausea, vomiting AOM episodes
Funding sources	"This work was funded by the National Health and Welfare Research and Development Program, Ottawa, Canada (grant 6606-2944-42). The sulfisoxazole was kindly provided by Hoffmann Laroche Canada Ltd."
Declarations of interest	No declaration was made.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). Baseline characteristics are not excessively similar. Plausible loss to follow-up reported. No implausible results.

	The number randomised to each group was not identical.

Risk of bias		
Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	The method used for sequence generation was not reported.
Allocation concealment (selection bias)	Unclear risk	No attempt to conceal allocation was reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Blinding of participants and personnel is not reported. There is a strong possibility that participants and personnel could identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	High risk	The only outcome reported to have been conducted blind to treatment allocation was tympanometry, “tympanometry was conducted only at 18 months to keep the audiologist “blind” to treatment group”. However, the other outcomes of episodes of AOM and some adverse events, such as rash and nausea, are more likely to be influenced by lack of blinding. Thus, some outcomes are at low risk of detection bias and others are at high risk, giving an overall rating of high.
Incomplete outcome data (attrition bias) All outcomes	Low risk	8 of 68 (12%) subjects in the VT group and 6 of 71 (8%) were lost to follow-up. Reasons for loss to follow-up were reported as subjects moving out of town and parental refusal to attend follow-up appointments.
Selective reporting (reporting bias)	Unclear risk	No protocol or trial registration was found. All outcomes specified in the published paper were reported.
Other bias	High risk	1. The first 10 surgical subjects received a different VT to subsequent subjects. A different VT was used for later participants as it was reported that these VT were "more effective in managing hearing loss". The authors do not consider the effect of the use of different VT on outcomes. 2. 31 of 65 (48%) medically treated participants were retreated with VT and 6 of 60 (10%) were retreated with sulfonamide. Analysis was according to the ITT principle.

D'Eredita 2006

Study characteristics	
Methods	2 arm, parallel group, non-blinded, single centre, non-blinded RCT with randomisation by child and 12 months follow-up.
Participants	**Location:** Italy, single centre **Setting of recruitment and treatment:** Division of Paediatric Otolaryngology, in a tertiary paediatric care institution. **Study dates:** January 2001 to January 2003 **Sample size:** • **Number randomised:** 30 [15 in VT group, 15 in laser myringotomy group] • **Number completed:** 30 [15 in VT group, 15 in laser myringotomy group] **Participant (baseline) characteristics:** Age (years): • Ventilation tubes (VT): 3.6 (range 2 to 6); • Laser myringotomy (LM): 3.8 (range 2 to 6) Gender: • VT M 8/15 (53%) F 7/15 (47%); • LM 8/15 (53%) F 7/15 (47%) **Inclusion criteria:** • OME for at least 3 months duration **Exclusion criteria:** • a history of prior middle ear surgery or pressure equalising tube insertion

	• Down or other syndrome involving the head and neck • cleft palate or previous pharyngeal surgery • mental retardation or other known cognitive or psychiatric disorder
Interventions	**VT group:** cold myringotomy, middle ear secretions were suctioned and a Teflon Shah tube inserted. **Laser myringotomy:** laser myringotomy using diode laser, then middle-ear secretions suctioned. Laser settings were 2 W power, 0.5 s pulse duration, with five pulses in the contact mode used with 600 mm thick fibre which tapers to a 300 mm tip. **Use of additional interventions:** Following VT or LM, "middle ear secretions were suctioned. Ofloxacin 0.3% otic solution (Floxin otic1, Daiichi Pharmaceutical Corp., Montvale, NJ) was then instilled in each ear, and was prescribed for use at home thrice daily for 5 days."
Outcomes	Hearing returned to normal • no definition of normal hearing was provided Persistent perforation Otorrhoea
Funding sources	Not reported
Declarations of interest	Not reported
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). Baseline characteristics show identical numbers of males/females. No loss to follow-up was reported. Hearing was assessed as normal in all children at follow-up, which may be implausible. The number randomised to each group was identical, and no information on how randomisation was performed.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Quote: "Thirty children ... with OME for at least 3 months duration were randomized into study (CDLM) and control (M&T) groups." No details provided.
Allocation concealment (selection bias)	Unclear risk	No details provided.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Blinding of participants and personnel is not reported. There is a strong possibility that participants and personnel could identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No details provided.
Incomplete outcome data (attrition bias) All outcomes	Low risk	The only missing data seems to be one of 60 parent-completed questionnaires. No children were lost to follow-up.
Selective reporting (reporting bias)	High risk	"Patients were scheduled for post-operative office evaluation at day 10, 20, 30, 40, 60 and 80, and then at month 3, 4, 5, 6, 8, and 12. During each visit, myringotomy patency and tube status were assessed ….. All patients underwent a post-operative age-appropriate audiometric evaluation with tympanometry at month 6, and then again at 1-year follow-up." No protocol is available. The main outcome of middle ear ventilation is presented graphically. However, data presented in text are sparse. Little outcome data is presented for tympanometry and audiometric testing at six and 12 months.
Other bias	Unclear risk	

		No details given as to how potential participants were identified for the study. The instructions given to parents in completing the questionnaire, the validity of the questionnaire and the reliability of outcome assessments were not reported. The risk of detection bias is therefore unclear.

Dempster 1993

Study characteristics	
Methods	Single centre RCT with 11 months follow-up. Randomisation by child for adenoidectomy, and subsequently one ear was randomly selected to receive a ventilation tube. Data of relevance for this review is the comparison of unilateral ventilation tube versus no treatment in ears of the same individual (either with no additional surgery, or with a background of adenoidectomy)
Participants	**Location:** UK, single centre **Setting of recruitment and treatment:** Paediatric hospital clinic in Glasgow. **Study dates:** August 1986 to February 1989 **Sample size:** • **Number randomised:** 78 (number allocated to each group not reported) • **Number completed:** 72 (37 with adenoidectomy, 35 without adenoidectomy) **Participant (baseline) characteristics:** Age, years, SD (range): • Adenoidectomy (with and without VT) = 5.9 +/- 1.4 (4 to 9) • No adenoidectomy (with and without VT) = 5.7 +/- 1.2 (4 to 9) Gender • Adenoidectomy (with and without VT) = 17 males (46%) : 20 females (54%) • No adenoidectomy (with and without VT) = 23 males (66%) : 12 females (34%) **Inclusion criteria:** • Children aged between three and a half and 12 years • Otoscopic evidence of bilateral otitis media with effusion that satisfied the following criteria on two assessments, 12 weeks apart: ◦ (a) pure tone air conduction thresholds average over 0.5, 1 and 2 kHz of ≥25 db HL ◦ (b) an air-bone gap over 0.5, 1 and 2 kHz of ≥15 dB ◦ (c) Type B tympanogram **Exclusion criteria:** • previous adenoidectomy or aural surgery • additional symptoms requiring surgical intervention, e.g. recurrent sore throat • cleft palate.
Interventions	**Intervention and comparisons** Ventilation tube insertion: • A unilateral Shah grommet was inserted following a radial myringotomy with aspiration of fluid Control group: • The contralateral ear was not operated on. The comparison was made between the ears of the same individual (operated versus un-operated side). Note that half of the children in this trial also underwent adenoidectomy. For the purposes of this review we have displayed the data from children who underwent adenoidectomy separately to those who did not undergo adenoidectomy. However, the data have been pooled together, to show the overall effect of ventilation tubes (with or without adenoidectomy).
Outcomes	Proportion of ears with hearing returned to normal

	• defined by the study authors as <15dB HL, using air conduction thresholds from pure tone audiometry. Mean final hearing threshold (air conduction and air-bone gap) • pure tone air conduction thresholds and air-bone gap thresholds averaged over 0.5, 1 and 2 kHz Mean change in hearing threshold Proportion of ears with persistence of OME • Assessed using both otoscopy and tympanometry. Adverse events: • Proportion of ears with perforation/retraction • Proportion of ears with tympanosclerosis • Proportion of ears with tube not *in situ*
Funding sources	Not reported.
Declarations of interest	No declaration is made.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). No excessive similarities in baseline characteristics. Plausible loss to follow-up reported. No implausible results. The number randomised to each group was not reported.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No details provided on how allocation sequence was generated.
Allocation concealment (selection bias)	Low risk	"These 78 children were then admitted to hospital within ten days and randomly allocated by a serially numbered envelope system..."
Blinding of participants and personnel (performance bias) All outcomes	High risk	No information provided on blinding of participants and personnel. There is a strong possibility that participants and personnel could identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	“At six and 12 months post-surgery, the presence or absence of otitis media in the non-grommeted ear was record by the validated otoscopist who was blind as to whether adenoidectomy had been performed and by tympanometry.” There was no report of blinding for either tympanometric or audiometric assessment. The outcomes are not sufficiently objective to discount the possibility of ascertainment bias.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	"Six children defaulted either at the six or 12 month assessment visits, leaving 72 (92 per cent) children with complete clinical, audiometric and tympanometric data for the pre-operative and these post-operative visits." Six of the 78 (8%) randomised children were lost to follow-up. The distribution of those six across groups is not reported. Precise reasons for losses to follow-up were not reported. It is therefore difficult to judge the potential for attrition bias.
Selective reporting (reporting bias)	Unclear risk	No protocol or trial registration was found. The published paper reports all expected outcomes.
Other bias	Unclear risk	It is unclear whether (for VT versus no treatment) comparisons were made within each individual child. The data are presented as if comparisons were made at whole trial arm level, as in a parallel group trial. There could therefore be a unit of analysis error, which could result in spuriously wide confidence intervals.

Elkholy 2021

Study characteristics	
Methods	Single centre parallel group RCT with 1 year of follow-up.
Participants	**Location:** Egypt, single centre **Setting of recruitment and treatment:** ENT and paediatric outpatient clinics at Al-Azhar University Hospital, Cairo. **Study dates:** September 2018 to March 2020 **Sample size:** • **Number randomised:** 40 • **Number completed:** 40 **Participant (baseline) characteristics:** Age, years (SD): • Ventilation tubes plus adenoidectomy: 7.3 years (1.90) • Adenoidectomy alone: 6.1 years (1.2) Sex • Ventilation tubes plus adenoidectomy: 8 males: 12 female • Adenoidectomy alone: 10 male: 10 female **Inclusion criteria:** • children with OME and adenoid hypertrophy, aged 5 to 15 years old; • persistent or recurrent OM despite proper medical treatment for 3–6 months. **Exclusion criteria:** • Children with naso-facial malformation, cleft palate or allergic rhinitis • A history of adenoid operation or ventilation tube insertion • Any other ear problem
Interventions	**Intervention:** Ventilation tube insertion (unclear if one or both ears) and adenoidectomy. N = 20. **Comparator:** Adenoidectomy alone. N = 20.
Outcomes	Persistence of OME at 2 weeks follow-up.
Funding sources	Not stated.
Declarations of interest	The authors state that they have no conflict of interest.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration was not identified. Baseline characteristics are not excessively similar. No reason is given for full follow-up No implausible results were identified. The number randomised to each group was identical, and there is no description of block randomisation.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	High risk	Quote: “Included children were randomly divided into two groups based on the consecutive number of enrollments those with odd number were included into group A while those of even number were included in group B” Comment: Quasi-randomised allocation.
Allocation concealment (selection bias)	High risk	Quote: “Included children were randomly divided into two groups based on the consecutive number of enrollments those with odd number were included into group A while those of even number were included in group B”

		Comment: Quasi-randomised allocation, allowing group allocation to be predicted.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Participants and study personnel would have been aware of the group allocation. No blinding was used.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No indication is given that outcome assessors were blinded. Outcomes were assessed by study personnel, therefore we assume they were aware of the group allocation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Full follow-up is reported.
Selective reporting (reporting bias)	Unclear risk	No protocol was available to assess the intended reporting plan.
Other bias	High risk	Data were only available after 2 weeks of follow-up, which is too short to fully assess the benefit of this intervention. Data from later time points were incompletely reported, precluding their inclusion in the review.

Gates 1989

Study characteristics	
Methods	Parallel group, four-arm, multicentre RCT with 2 years duration of follow-up. Randomisation by child. This study included a comparison of ventilation tubes, myringotomy and adenoidectomy. For the purposes of analysis we have compared children who received ventilation tubes with those who received myringotomy, and also compared children who received ventilation tubes plus adenoidectomy to those who received myringotomy plus adenoidectomy.
Participants	**Location:** USA, multicentre **Setting of recruitment and treatment:** Hospital-based otitis media study centre in the US. Inpatient and outpatient management. Fourteen participating otolaryngologists in five hospitals. **Study dates:** Not reported **Sample size:** • **Number randomised:** 578 • **Number completed:** 389 **Participant (baseline) characteristics:** Age, years • VT alone: ◦ 89/129 (69%) aged 4 to 6.5 years ◦ 40/129 (31%) aged 6.5 to 8 years • VT plus adenoidectomy: ◦ 92/125 (74%) aged 4 to 6.5 years ◦ 33/125 (26%) aged 6.5 to 8 years • Myringotomy alone: ◦ 74/107 (69%) aged 4 to 6.5 years ◦ 33/107 (31%) aged 6.5 to 8 years • Adenoidectomy plus myringotomy: ◦ 95/130 (73%) aged 4 to 6.5 years ◦ 35/130 (27%) aged 6.5 to 8 years Gender • VT alone: 89 (59%) male: 61 (41%) female • VT plus adenoidectomy: 88 (59%) male: 62 (41%) female • Myringotomy alone: 76 (60%) male: 51 (40%) female • Adenoidectomy plus myringotomy: 90 (60%) male: 61 (40%) female

	Hearing loss at baseline
	• VT alone:
	  ◦ better ear 23.13 db HL
	  ◦ worse ear 34.41 db HL
	• VT plus adenoidectomy:
	  ◦ better ear 23.93 db HL
	  ◦ worse ear 27.05 db HL
	• Myringotomy alone:
	  ◦ better ear 24.49 db HL
	  ◦ worse ear 37.26 db HL
	• Adenoidectomy plus myringotomy:
	  ◦ better ear 24.86 db HL
	  ◦ worse ear 26.12 db HL
	Inclusion criteria:
	• Children age 4 to 8
	• otolaryngologist-confirmed chronic middle ear effusion, persisting 60 days after a course of 10 days erythromycin 50mg/kg and sulfisoxazole 150mg/kg, and 30 days of pseudoephedrine hydrochloride 4mg/kg.
	Exclusion criteria:
	• History of prior tonsil or adenoid operations
	• VT placement (within 2 years)
	• cleft palate
	• major chronic illness, required daily medication (other than anti-allergy therapy)
	• other otologic diagnoses, advanced or irreversible structural changes of the tympanum (such as cholesteatoma, permanent perforation or atelectasis).

Interventions	**Intervention and comparisons** Bilateral myringotomy • Both TMs were opened regardless of operative otoscopic findings, unless one ear had been perfectly normal on all preoperative otoscopic examinations. • n=127 VT • Shepard type with 1.1mm internal opening. Both TMs were opened regardless of operative otoscopic findings, unless one ear had been perfectly normal on all preoperative otoscopic examinations. • n=150 Adenoidectomy and myringotomy • Adenoidectomy by curettage with mirror plus myringotomy as above • n=151 Adenoidectomy and VT • Adenoidectomy and ventilation tube insertion tube as above. • n=150

Outcomes	**Primary outcomes relevant to this review:**
	• **Hearing**
	  ◦ Only assessed as the proportion of time with any hearing loss. The number of visits in which a child had a hearing threshold of ≥20 dB, (using the three-frequency, pure-tone average) was divided by the number of visits made, and weighted for the number of visits made. This proportion was determined for each child and averaged for each group. These data were not included in the review.
	• **Disease-specific quality of life**
	  ◦ Not reported

	• **Adverse event**
	  ◦ Haemorrhage
	Secondary outcomes relevant to this review:
	• **Presence/persistence of OME: proportion of children with persistence of OME**
	  ◦ Persistence was determined using an algorithm based on otoscopy and tympanometry. Also reported as the proportion of time with an effusion.
	• **Other adverse effects**
	  ◦ Not reported

Funding sources	Supported by National Institutes of Health/National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) contract NO1 NS 02328 and a grant in kind from Ross Laboratories.
Declarations of interest	None reported.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). Baseline characteristics are not excessively similar. Plausible loss to follow-up reported. No implausible results. The number randomised to each group was not identical.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"If informed consent was given, the child was assigned randomly by the project statistician, using tables of random numbers, to one of four groups...". This method would be expected to produce an adequate balance of prognostic factors across groups. However, two issues were reported, that might have interfered with the balance produced by randomisation: (1) parents of children were free after randomisation to choose an alternative treatment; and (2) there were fewer patients in group 1 because entry was stopped early at the request of the Safety and Data Monitoring Board. However, reported patient characteristics were adequately balanced across groups, suggesting that randomisation was adequate.
Allocation concealment (selection bias)	Low risk	"If informed consent was given, the child was assigned randomly by the project statistician, using tables of random numbers, to one of four groups...". As allocation was undertaken by the statistician, allocation was probably concealed.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Parents of children were informed of treatment allocation. Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Despite otoscopists being blind to treatment allocation and outcome data, treatment allocation would be obvious in instances when a VT is visible. Otoscopic assessments have a degree of subjectivity.
Incomplete outcome data (attrition bias) All outcomes	High risk	Despite losses to follow-up being of similar proportions across groups, and despite the characteristics of those losses being similar to those who were not lost to follow-up, the very high attrition rate of 189/578 (33%) constitutes a major loss of data, exceeding the effect size for outcomes relating to persistence of effusion.
Selective reporting (reporting bias)	Low risk	No protocol was available, but pre-specified outcomes were reported.
Other bias	High risk	The parents of 27 of the 491 randomised children (5.5%) chose a treatment other than that to which their child was randomised. Of 491 children, 240 (49%) received medical retreatment for chronic effusion. Of

		491 children, 109 (22%) met the criteria for surgical retreatment. Given the number of children receiving retreatment, there is the strong possibility of contamination within the trial.

Koopman 2004

Study characteristics	
Methods	2 arm, multicentre parallel-group RCT, with randomisation by ear and 6 month follow-up.
Participants	**Location:** Netherlands, 7 sites **Setting of recruitment and treatment:** paediatric hospital **Study dates:** July 1999 to September 2001 **Sample size:** 208 children (416 ears) • **Number randomised:** 208 ears in laser myringotomy, 208 ears in VT • **Number completed:** 153 ears in laser myringotomy, 153 ears in VT **Participant (baseline) characteristics:** Age (mean (SD) years): 4.2 (2.3) (for all 208 children) Gender: M 108/208 (52%) F 100/208 (48%) Duration of disease: 6 months (range 3 to 12 months) Treatment used before trial entry: Adenoidectomy, tonsillectomy, and grommets in 24.5%, 11.1% and 23.6% of patients, respectively. **Inclusion criteria:** • children aged less than 11 years • impaired hearing noticed by parents during at least 3 successive months • bilateral OME. **Exclusion criteria:** • unilateral OME • poorly cooperative children • clinically admitted patients • asymmetric perceptive HL • previously operated ears with other than myringotomy or ventilation tubes.
Interventions	All participants had one intervention in each ear. **Laser myringotomy:** performed with a Sharplan CO2-flashscanner laser using a handhold device and video screen (ESC Sharplan Medical Systems, Tel Aviv, Israel). The power setting varied from 7 to 20 W, and the diameter of the circular perforation varied from 1.8 to 2.6 mm, with an aim for the largest diameter as possible (2.6 mm in 159 of 208 patients). The laser myringotomy was performed in the anteroinferior part of the tympanic membrane without aspiration of fluid. **Ventilation tube:** inserted using cold-knife myringotomy, A ventilation tube with a 1.1 mm internal diameter (Donaldson) was used (94%). In case of OME with atelectasis of the middle ear, a Goode-T Tube (6%) was inserted in the anteroinferior part of the tympanic membrane. **Use of additional interventions:** Adenoidectomy in combination with tonsillectomy was performed in 16 children. Otorrhoea persisting for more than 1 week was treated by ear drops consisting of either dexamethasone/framycetine/gramicidin or ofloxacin, depending on the culture, whereas otorrhoea with fever was treated with oral antibiotics only (amoxicillin). During administration of medication, the child was seen weekly until recovery.
Outcomes	Proportion of children with persistence of OME Adverse events • otorrhoea • otalgia
Funding sources	The Sophia Foundation For Medical Research and The Revolving Fund Sophia Children's Hospital, Erasmus Medical Centre, Rotterdam, Theia Foundation, and Silver Cross Company.

Declarations of interest	"The authors declare that there is no conflict of interest of any kind in this study"
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). Baseline characteristics are not relevant (split-body trial) Plausible loss to follow-up reported. No implausible results. The number randomised to each group was identical as this was a split-body trial.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	'Assignment of the side for laser myringotomy or tube insertion was made randomly by computer-generated lists in balanced blocks of six to assure an even distribution of surgical procedure for left and right ears.'
Allocation concealment (selection bias)	Unclear risk	No method of allocation concealment was reported.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Surgeons could not be blinded. There is a strong possibility that personnel could identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	No information provided on blinding of outcome assessors.
Incomplete outcome data (attrition bias) All outcomes	High risk	The rate of loss to follow-up was high: 'A total of 55 (26%) children quit the study (41 lost to follow-up, 14 failures). The frequency of control visits was the main reason for discontinuation of follow-up.' There was no detailed account of reasons for losses to follow-up. The proportion of missing outcomes (26%) compared with observed event risk (e.g. proportion effusion free after laser myringotomy at 3 months 37.1%) could be enough to induce clinically relevant bias in intervention effect estimate.
Selective reporting (reporting bias)	High risk	One or more outcomes of interest in the review (e.g., otorrhoea and perforation) are reported incompletely, and thus cannot be entered in a meta-analysis.
Other bias	Unclear risk	A follow-up period of six months may be too short to assess a clinically meaningful outcome of persistence of OME.

Maw 1983

Study characteristics	
Methods	Randomised, parallel group, single blind controlled trial of adenotonsillectomy or adenoidectomy or no pharyngeal surgery, with three years of follow-up. Split-body randomisation was used to place a VT in one ear of each participant. For the purposes of this review we have included data comparing the ear with the ventilation tube to the un-operated, contralateral ear in the same participant. Only participants who did not receive additional surgery were included in this analysis.
Participants	**Location:** UK, single centre **Setting of recruitment and treatment:** UK inpatient and ENT outpatient setting in Bristol. **Study dates:** Recruitment started in July 1979. End-date not reported. **Sample size:** Note that this is the sample size for relevant arms included in this review, not the total sample size for the whole trial (N = 192). • **Number randomised:** 56 • **Number completed:** 47

	Participant (baseline) characteristics: Age, years, SD (range): 5.31 years (SD 1.22) Gender: 32 males (57%), 24 females (43%) **Inclusion criteria:** • persistent subjective hearing difficulty; • pneumatic otoscopic confirmation of bilateral effusions; • symmetrical audiometric hearing loss, in excess of 25 dB at one or more frequencies; • impedance measurements not showing a peak A type of curve. **Exclusion criteria:** • resolution of fluid over subsequent 12 weeks • medical grounds, mostly because of upper airway obstruction from gross adenoidal hyperplasia (often with sleep apnoea) • refused random allocation • asymmetrical hearing loss or because a super added sensorineural loss was suspected • preoperative follow-up was less than three months
Interventions	**Intervention and comparisons** Ventilation tube insertion: • One ear of all children was treated at random with ventilation tube insertion. Control: • The contralateral ear was left un-operated. Background treatments: No additional surgery was used for participants included in this review. Other participants in the study did undergo adenoidectomy or adenotonsillectomy.
Outcomes	Final hearing threshold (operated and un-operated ear).
Funding sources	Not reported.
Declarations of interest	Not reported.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). No excessive similarities in baseline characteristics. Plausible loss to follow-up reported. No implausible results. The number randomised to each group was similar but not identical.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"From tables of random numbers, the children were allocated as follows: adenotonsillectomy 47; adenoidectomy 47; no-surgery 56."
Allocation concealment (selection bias)	Unclear risk	The method of concealment is not described.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Surgeons could not be blinded. There is a strong possibility that personnel could identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	High risk	"The accuracy of A. R. M. (the clinical investigator) in otoscopic diagnosis has been assessed and reported previously." The lead researcher undertook the pneumatic otoscopy. Blinding of audiometric and tympanometric assessments was not reported and therefore assessments are unlikely to be blinded. Audiometry is open to subjective assessment.

Incomplete outcome data (attrition bias) All outcomes	Unclear risk	The attrition rate was similar in each group of interest (24% and 23% at one year, and 53% and 52% at three years, in the adenoidectomy plus unilateral VT group and the unilateral VT group, respectively). The reasons for attrition were largely unreported and could have been related to the outcomes of interest.
Selective reporting (reporting bias)	Low risk	No published protocol has been found, but it appears that all pre-specified outcomes are reported.
Other bias	Low risk	None identified.

Maw 1999

Study characteristics	
Methods	Parallel group, single-centre 2-arm RCT with up to 7 years of follow-up. Randomisation by child.
Participants	**Location:** UK, single centre **Setting of recruitment and treatment:** Paediatric hospital clinic in Bristol. **Study dates:** November 1993 to January 1996 **Sample size:** • **Number randomised:** 182 (92 to ventilation tubes, 90 to watchful waiting) • **Number completed:** 156 to 18 months (83 to ventilation tubes, 73 to watchful waiting) **Participant (baseline) characteristics:** Age, years, SD (range): • VT = 2.96 (0.84) • Watchful waiting = 2.93 (0.87) **Inclusion criteria:** • date of birth between April 1, 1991 and Dec 31, 1992 (aged 9 months to approximately 4.5 years) • confirmation of bilateral OME by otoscopy and tympanometry (bilateral type B or C2 tympanograms and hearing loss of 25-70dB); assessment of hearing loss • disruptions to speech, language, learning, or behaviour. **Exclusion criteria:** • cleft palate • syndromes such as Down's, Hunter's, or Hurler's
Interventions	**Intervention and comparisons** **Ventilation tubes:** Surgery was by insertion of bilateral ventilation tubes. In children with clinical evidence of nasal obstruction because of adenoid enlargement, adenoidectomy was also done. In the early-surgery group, if hearing difficulty returned, otoscopy showed recurrence of effusions, with type B or C2 tympanograms during follow-up, tube reinsertion would be performed, if desired, within 6 weeks. **Watchful waiting** Participants were advised that - if the need for an operation was recognised at the 9-month assessment - surgery would be done within 6 weeks of that date. Approximately 21% of participants received surgery before 9 months of follow-up. By 18 months, only 15% of participants in this group had not been listed for, or already received surgery.
Outcomes	Final hearing threshold (right ear, left ear, best ear, worst ear) • assessed with pure tone audiometry at 4000Hz. Proportion of children with persistence of OME by otoscopy and tympanometry in one or both ears, and in the best ear Receptive language skills (Reynell Language Scales) Speech development (Reynell Language Scales) Cognitive development (Griffiths Mental Development Scales) Psychological outcomes (Goodman)

	Listening skills
Funding sources	"The trial was funded by the South and West NHS Research and Development Directorate."
Declarations of interest	No declaration is made.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). No excessive similarities in baseline characteristics. Plausible loss to follow-up reported. No implausible results. The number randomised to each group was not identical.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomisation was performed using a random number table to generate numbers in an office distant from the hospital".
Allocation concealment (selection bias)	Low risk	"Numbers were placed in sealed envelopes".
Blinding of participants and personnel (performance bias) All outcomes	High risk	Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	"Tympanometry and hearing tests at randomisation and 9-month and 18-month follow-up visits were done by audiological scientists or technicians who were masked to the children's treatment status". "Audiological Scientists, Reynell Language and Griffith Mental Development scale testers were blind to allocation of treatment group. The Richman Behaviour Checklist was completed by parents." Therefore there is the potential for psychological outcomes (those assessed using the Richman Behaviour Checklist and behaviour total scores as reported by parents) to have been influenced by lack of blinding.
Incomplete outcome data (attrition bias) All outcomes	High risk	Twenty participants were lost to follow-up (4 in the VT group, 16 in the watchful waiting group) by final follow-up when participants were 7 years of age. It is unclear whether it is the same participants who were lost to follow-up at each follow-up period, as the number of participants for whom outcome data are available fluctuates throughout the years. There is an imbalance in numbers of missing data across intervention groups, and there is likely to be imbalance in reasons for missing data across intervention groups, for example, the authors note that "mothers of lower educational achievement provided complete data on these factors less often than other mothers" (Hall 2009, p 17). Additionally, authors note that "the validity of the results needs to be considered in the light of a number of factors […] loss to follow-up – although relatively low (9% in the early surgery and 18% in the watchful waiting group) – could introduce some degree of bias".
Selective reporting (reporting bias)	Unclear risk	No published protocol or trial registrations were found. For the outcome mean final hearing threshold for best ears at 9 months follow-up, two different sets of data at the same follow-up time point are presented in Maw 1999 vs Maw 2000. Authors note data were available for more children in Maw 2000 than in Maw 1999 for some outcomes, but it is unclear why this is the case.
Other bias	Low risk	The study appears to be free of other sources of bias.

Paradise 2007

Study characteristics	
Methods	Multicentre RCT with 11 years of follow-up. Randomisation by child.
Participants	**Location:** multiple sites in the USA **Setting of recruitment and treatment:** recruited from 2 urban hospitals, 2 small-town/rural and 4 suburban private paediatric practices

	Study dates: Recruitment from May 1991 to December 1995 **Sample size:** • **Number randomised:** 429 (216 to early treatment, 213 to watchful waiting) • **Number completed:** 391 (195 from early treatment group, 196 from watchful waiting group) **Participant (baseline) characteristics:** Age, months: mean 15 months for the whole cohort (median 14 months) Gender: • Early treatment group: 115 male (56.4%) : 89 female (43.6%) • Watchful waiting group: 112 (58%) male: 81 (42%) female **Inclusion criteria:** • OME beginning from the age of 2 months and within the first 3 years of life • middle-ear effusion that appeared substantial in degree and that persisted, despite treatment with antimicrobial drugs, for 90 days in the case of bilateral effusion or 135 days in the case of unilateral effusion • children with intermittent bilateral or unilateral middle-ear effusion for specified proportions of longer periods were also eligible. For example, a child would be eligible if he or she had had bilateral effusion for 67 percent of the preceding 180-day period **Exclusion criteria:** • birth weight less than 5 lb (2268 g) • small for gestational age • history of neonatal asphyxia or other serious illness • major congenital abnormality or chronic illness • multiple birth • sibling enroled in the study • in foster care or adopted before enrolment • mother dead, seriously ill, a known drug or alcohol abuser before enrolment • mother judged by study personnel to be unable to give informed consent or adhere to the study protocol • mother less than 18 years of age • English not the only household language (from ClinicialTrials.gov)
Interventions	**Intervention and comparisons** **Early treatment (VT)** Children were scheduled to have ventilation tubes inserted as soon as possible (n=216 randomised; 195 completed follow-up and 164 had received ventilation tubes by the age of 9-11 years). **Watchful waiting/Late treatment (VT)** Children were scheduled to have ventilation tubes after a six-month delay (if bilateral effusion persisted) or after a 9-month delay (if unilateral effusion persisted) (n=213 randomised; 196 completed follow-up and 88 had received ventilation tubes by the age of 9-11 years).
Outcomes	Proportion of children with normal hearing returned to normal • Defined by the authors as ≤15 db HL Mean final hearing threshold (left ear, right ear) Persistence of OME (none, unilateral, bilateral, indeterminate) Adverse event: • persistent perforation • tympanosclerosis • fibrosis • segmental atrophy Receptive language skills

	Speech development Cognitive development Psychological development Listening skills Parental stress
Funding sources	"Supported by grants from the National Institute of Child Health and Human Development and the Agency for Healthcare Research and Quality (HD26026 and HD42080), from the University of Pittsburgh Competitive Medical Research Fund, and from the Children's Hospital of Pittsburgh Research Advisory Committee and by gifts from GlaxoSmithKline and Pfizer."
Declarations of interest	None declared.
Notes	No retraction notices identified. Prospective registration not applicable (published before 2010). No excessive similarities in baseline characteristics. Plausible loss to follow-up reported. No implausible results. Block randomisation was used to ensure balanced allocation to the two groups.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Assignments were made by designated nonclinical staff members using separate, computer-generated lists of random numbers."
Allocation concealment (selection bias)	Unclear risk	"Assignments were made by designated nonclinical staff members using separate, computer-generated lists of random numbers." It is unclear the role these staff members played in the study and thus it is difficult to judge whether their knowledge of the sequence influenced allocation and had a possible effect on outcomes.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result. The parents of the child would know the allocation and it might affect their behavior or decision to use adjunctive treatments.
Blinding of outcome assessment (detection bias) All outcomes	Low risk	• Examiners and analysts carrying out developmental tests were unaware of the children's medical histories and treatment-group assignments at follow-up when participants were 9-11 years of age, but no information about blinding of other outcome assessors, such as audiologists, is provided. • Examiners, transcriptionists, and analysts were blinded to the children's health histories including receipt of tympanostomy tubes at follow-up when participants were 6 years of age, but no information about blinding of other outcome assessors, such as audiologists, is provided. • All otomicroscopic examinations were conducted by a pediatric otolaryngologist who was unaware of children's history and study group assignment, and audiologists were unaware of children's otoscopic diagnoses at follow-up when participants were 5 years of age.
Incomplete outcome data (attrition bias) All outcomes	Low risk	At age 3, 206 of 216 (95%) who had early treatment underwent developmental tests and 196 of 213 (92%) who had late treatment underwent developmental tests. At age 4, 204 of 216 (94%) who had early treatment underwent developmental tests and 193 of 213 (91%) who had late treatment underwent developmental tests. No reasons are given for attrition/exclusion but low levels.
Selective reporting (reporting bias)	Low risk	There is a trial registration for study of 9-11 year olds. It appears that all pre-specified outcomes are reported for each time of assessment.
Other bias	Low risk	There does not appear to be any other sources of bias.

Popova 2010

Study characteristics	
Methods	Parallel group, single centre RCT with 12 month follow-up. Randomisation by child.
Participants	**Location:** Bulgaria, single centre

	Setting of recruitment and treatment: ENT department of University Hospital "Queen Jovanna", Sofia, Bulgaria **Study dates:** 2007-2009 **Sample size:** • **Number randomised:** 90 • **Number completed:** 78 **Participant (baseline) characteristics:** Age, years, SD: • Ventilation tubes: mean 60 months (SD 11.6) • Myringotomy: mean 61 months (SD 9.4) Gender • Ventilation tubes: 22 (52%) male: 20 (48%) female • Myringotomy: 20 (56%) male: 16 (44%) female Hearing threshold at baseline • Ventilation tubes: mean 31.4 db HL (SD 6.4) • Myringotomy: mean 32.3 db HL (SD 6.5) **Inclusion criteria:** • history of bilateral middle ear effusion for at least 3 months • conductive hearing loss greater than 20 dB **Exclusion criteria:** • previous myringotomy with or without insertion of ventilation tubes • previous adenoidectomy or tonsillectomy • history of ear surgery • cleft palate • Down's syndrome • congenital malformations of the ear • cholesteatoma or chronic mastoiditis • perforation of the tympanic membrane • conductive hearing loss attributed to destructive changes in the middle ear • sensorineural hearing loss
Interventions	**Intervention and comparisons** Adenoidectomy and VT • Adenoidectomy was performed using electrocautery, curette and St. Clair-Thomsen forceps. Tympanostomy tubes were inserted again in the inferior-posterior portion of pars tensa after an incision was made in this location and aspiration of the effusion was assured. All of the inserted ventilation tubes were fluoroplastic Donaldson grommets (Micromedics, Inc.) • n=42 Adenoidectomy and myringotomy • Adenoidectomy was performed using electrocautery, curette and St. Clair-Thomsen forceps whereas myringotomy consisted of a wide incision in the inferior-posterior portion of pars tensa followed by aspiration of the effusion. • n=36
Outcomes	Mean final hearing threshold Proportion of children with persistence of OME Adverse events: • tube occlusion • premature extrusion • otorrhoea

	Episodes of AOM
Funding sources	No details are given.
Declarations of interest	"Authors report no conflict of interest in the publication of the article"
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration was not identified (published in 2010). Baseline characteristics are not excessively similar. Plausible loss to follow-up reported. No implausible results. The number randomised to each group was not identical.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	No details are given.
Allocation concealment (selection bias)	Unclear risk	No details are given.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding of outcome assessment is reported so we assume no blinding and therefore a high risk of bias.
Incomplete outcome data (attrition bias) All outcomes	High risk	“Ninety patients with bilateral OME were enrolled initially in our study. Seventy-eight of them (156 ears) attended all of the appointed examinations during the whole follow-up period and remaining twelve were excluded.” Data are not available for these 12 participants, including which intervention they received. It is possible that the reason for missing data for these participants could be related to true outcome.
Selective reporting (reporting bias)	Unclear risk	No protocol or trial registration has been found. Authors did not clearly state the outcomes they would be assessing in the study.
Other bias	High risk	“All 5 patients with recurrence from the A+M group were treated conservatively with medications as described previously [9] and subsequently on one of them a tympanostomy tube was inserted, which followed to his exclusion from the A+M group.” Thus, this study appears to have adopted a per protocol analysis.

Rach 1991

Study characteristics	
Methods	Single centre RCT with 6 month follow-up, and additional follow-up of developmental outcomes for up to 4 years. Randomisation by child.
Participants	**Location:** Netherlands, single centre **Setting of recruitment and treatment:** Recruitment from GP surgeries, trial run from ENT clinic. **Study dates:** Not reported **Sample size:** • **Number randomised:** 43 (22 to ventilation tubes, 21 to control) • **Number completed:** 43 (22 to ventilation tubes, 21 to control) **Participant (baseline) characteristics:** Age, years, SD (range): • All participants aged 2-4 years Gender • Not reported

	Inclusion criteria: • Aged between 2 and 4 years • Bilateral flat tympanograms (Type B) at 2 screenings, 3 months apart • Dutch speaking **Exclusion criteria:** • congenital ear disorders (sensorineural loss) • defects in their speech-producing apparatus (e.g. cleft palate) neurological or serious visual disorders • emotional problems • mental health problems • chronic diseases • history of long-term (6 weeks or more) hospitalisation or chronic otorrhoea
Interventions	**Ventilation tubes** Standard (silicone ventilating tubes, Donaldson design). Insertion was performed bilaterally under GA in the antero-inferior quadrant of the tympanic membrane (n=22). **Comparator** No treatment (n=21). Note that some participants in this group may have undergone ventilation tube placement during the extended follow-up period (after a 6-month delay, and up to 7-8 years of age). Results until 6 months of follow-up are therefore included in Comparison 1 (VT versus no treatment) but results from extended follow-up are included in Comparison 2 (VT versus watchful waiting).
Outcomes	Proportion of ears with persistence of OME Adverse events: • tube extrusion Receptive language skills (Reynell) • reported as Z scores ([language score - mean score]/standard deviation), where higher scores reflect better skills Expressive language skills (Reynell) • reported as Z scores, as described above
Funding sources	This study was supported by a grant from the Dutch Prevention fund (no. 28-924).
Declarations of interest	None declared.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). No baseline characteristics are reported, therefore unable to assess. Loss to follow-up is unclear, but may be zero. No implausible results. Numbers allocated to each group are similar but not identical.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Randomized allocation was performed for the first five children entering the trial; each subsequent child was allocated to the treatment group which would lead to the smallest imbalance of the four determinants noted above. " As the process of minimisation si described this is low risk.
Allocation concealment (selection bias)	Unclear risk	No details on allocation concealment provided.
Blinding of participants and personnel (performance	High risk	Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result.

bias) All outcomes		
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	A rating of low risk of bias would be appropriate for grading the certainty of evidence for developmental test outcomes (receptive language skills and expressive language skills), because authors report that “All tests were performed and scored by one speech therapist, without previous knowledge of the child's history”. However, there was no report of blinding to treatment allocation for tympanometry.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Information on loss to follow-up is not reported, although the data reported indicate no loss to follow-up. However, authors note “The total group from whom two language tests could be obtained comprised 52 children”, indicating that only participants in the original prospective longitudinal study who had the necessary data at baseline and follow-up were included in this study. Therefore, there is potential that participants who were not available for follow-up were excluded from the study, although this is not reported in exclusion criteria (criteria only lists “not visiting the GP after referral” and “no referral by the GP to the ENT outpatient clinic” as exclusion reasons related to this issue). Authors do not give any further information, so it is difficult to judge the potential for attrition bias.
Selective reporting (reporting bias)	Low risk	There is no published protocol but it does not appear that selective reporting has occurred.
Other bias	Unclear risk	A follow-up period of 6 months is too short a time to show a real difference in language development, although other outcomes may be unaffected.

Rovers 2000

Study characteristics	
Methods	Multicentre randomised, controlled, parallel-group, open trial with 12 months of follow-up. Randomisation by child.

Participants	**Location:** Netherlands, multicentre study
	Setting of recruitment and treatment: 13 ENT hospital outpatient clinics in the Netherlands.
	Study dates: Recruitment from 1996 to 1998
	Sample size:
	• **Number randomised:**187
	• **Number completed:** 176
	Participant (baseline) characteristics:
	Age, years, SD (range):
	• Ventilation tubes: mean 19.5 months (SE 1.7)
	• Watchful waiting: mean 19.4 months (SE 1.9)
	Gender
	• Ventilation tubes: 55 males (59%) : 38 females (41%)
	• Watchful waiting: 55 males (59%) : 39 females (41%)
	Mean hearing threshold
	• Ventilation tubes:
	  ◦ best ear, mean 46.4dB
	  ◦ worst ear, mean 50.1dB
	• Watchful waiting:
	  ◦ best ear, mean 43.4dB
	  ◦ worst ear, mean 47.0dB
	Inclusion criteria:
	• Children who failed three successive hearing tests and were referred to an ENT outpatient clinic
	• Persistent bilateral OME confirmed by tympanometry and otoscopy, lasting for 4-6 months.

	Exclusion criteria: • Down syndrome • Sensorineural hearing loss • Cystic fibrosis • Asthma • Cleft palate
Interventions	**Ventilation tube insertion** Number randomised: 93. Number completed: 90. **Watchful waiting** Number randomised: 94. Number completed: 86. 10 children received treatment with ventilation tubes during the follow-up period (11.6%).
Outcomes	Change in hearing threshold • measured as the minimal response level using a portable visual reinforcement audiometry set. Reported as mean hearing thresholds in the better ear at 500, 1000, 2000 and 4000Hz. Difference in hearing improvement Persistence of OME Adverse events • otorrhoea • earache Receptive language skills (Reynell) • measured as the equivalent age - real age (higher scores indicate better development) Speech development (Schlichting) • measured as the equivalent age - real age (higher scores indicate better development) Erickson scale of parent-child interaction • Range from 1-7, higher scores = more interaction Generic HRQoL • using a modified version of the TAIQOL (TNO-AZL Infant Quality of Life) questionnaire. Rated on a 12 point scale - higher scores represent worse quality of life.
Funding sources	The Dutch Investigative Medicine Fund of the National Health Insurance Board.
Declarations of interest	None reported.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). Baseline characteristics are similar, but this is to be expected due to the balanced allocation procedure. Plausible loss to follow-up reported. No implausible results. Balanced allocation was reported.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	“To increase comparability at baseline, a balanced allocation procedure was employed with five balancing factors: sex, age, season at randomization, educational level of the mother, and hospital.” Minimisation was used.
Allocation concealment (selection bias)	Unclear risk	The method of concealment is not described.
Blinding of participants and personnel	High risk	Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result.

(performance bias) All outcomes		
Blinding of outcome assessment (detection bias) All outcomes	High risk	"During the trial, tympanometry and audiometry were performed by experienced audiologists (who were not blinded to the assignment of a child)." Some outcomes are likely to be influenced by lack of blinding.
Incomplete outcome data (attrition bias) All outcomes	High risk	Follow-up of 176/187 (94%) which is a high percentage, however eight were lost from the WW group and only three from the VT group. Furthermore 10 from the WW group went on to have VT.
Selective reporting (reporting bias)	Unclear risk	No protocol was available for comparison.
Other bias	Low risk	No protocol was available, but all pre-specified outcomes were reported.

Ruckley 1988

Study characteristics	
Methods	2 arm, parallel-group, single centre RCT, with randomisation by ear and 3 months follow-up
Participants	**Location:** Scotland, single centre **Setting of recruitment and treatment:** hospital **Study dates:** not reported **Sample size:** 40 children (80 ears) • **Number randomised:** *[40 in intervention, 40 in comparison]* • **Number completed:** *[36 in intervention, 36 in comparison]* **Participant (baseline) characteristics:** Age: 5 years 10 months (range 4 to 9 years) Gender: M 23/40 (58%) F 17/40 (42%) Duration of disease: >/= 3 months Baseline hearing loss (measured as the mean air-bone gap for the frequencies 0.25, 0.5, 1, 2 and 4 KHz): VT 21.4 dB (SD 6.5) thermal myringotomy group 21.0 dB (SD 6.6). **Inclusion criteria:** • first presentation with OME • bilateral OME for at least 3 months confirmed by audiometry, tympanometry and otoscopy. **Exclusion criteria:** not reported
Interventions	All participants received one intervention in each ear. **Ventilation tube:** myringotomy, with a conventional myringotomy knife, followed by aspiration of fluid and insertion of a Shepard grommet. **Thermal myringotomy:** using the Xomed thermovent device, followed by fluid aspiration. **Use of additional interventions:** all participants received adenoidectomy
Outcomes	**Primary outcome:** hearing assessed using air conduction and bone conduction **Secondary outcomes:** Appearance of tympanic membranes, patency of VT and thermal perforation, any otological symptoms, recurrence of middle ear fluid
Funding sources	Not reported.
Declarations of interest	Not reported.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). Baseline characteristics are not relevant (this is a split body trial) Plausible loss to follow-up reported. No implausible results.

	The number randomised to each group was identical, as this was a split-body trial.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	"Immediately prior to surgery a coin was spun in order to determine in a random fashion which ear was to be treated by thermal myringotomy."
Allocation concealment (selection bias)	Low risk	The need for allocation is obviated by using a simple method of randomisation at the point of intervention.
Blinding of participants and personnel (performance bias) All outcomes	Unclear risk	Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No blinding is reported and the authors do not clearly state who undertook outcome assessments. Otoscopy is sufficiently subjective for there to be a high risk of detection bias.
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	"Of the 40 children who entered the study complete results were obtained in 36. Four children failed to attend for regular post-operative review and were not included in the final results." As this study randomised by ear, loss of outcome data was equal for each intervention group. We do not know if the reasons for loss to follow-up were due to the intervention.
Selective reporting (reporting bias)	High risk	A study protocol is not available. One or more outcomes of interest in the review e.g. otalgia are reported incompletely.
Other bias	High risk	A follow-up period of 3 months is too short a time to assess the effect of the intervention.

Sujatha 2015

Study characteristics	
Methods	Randomised, parallel-group, open trial with 12 months of follow-up. Randomisation at the level of the individual child.
Participants	**Location:** India, single centre **Setting of recruitment and treatment:** Tertiary care hospital in Kerala. **Study dates:** January 2013 to December 2013 **Sample size:** • **Number randomised:** 50 [25 in VT plus adenoidectomy group, 25 in myringotomy plus adenoidectomy group] • **Number completed:** 50 [25 in VT plus adenoidectomy group, 25 in myringotomy plus adenoidectomy group] **Participant (baseline) characteristics:** Age (years): mean age 5.8 years (SD 1.8) Gender: 22 males (44%) and 28 females (56%) **Inclusion criteria:** • Age above 3 and below 10. • Children suffering from OME as diagnosed by impedance audiometry (Tympanometry), pure tone audiogram and pneumatic otoscopy. PTA air-bone gap should be at least 25db. • They should have taken medicines for OME (Steroid nasal spray 200microns/day in two divided doses, systemic decongestants and antihistamines) at least for 12 weeks but without any clinical benefit. • All children having associated adenoid hypertrophy (grade 3 or more) • Willing for randomisation into two groups and getting treatment specified in each group. **Exclusion criteria:** • Child known to have allergic rhinitis/taking medication for allergy/ bronchial asthma. • OME caused by any reason other than adenoid hypertrophy.

	• Not willing for randomisation and treatment strategy. • Children with cleft palate even if repaired. • Children with bifid uvula, Down/Turner syndrome. • Child having sensorineural hearing loss.
Interventions	Ventilation tube group: • adenoidectomy, myringotomy and ventilation tube insertion bilaterally. Shepard type ventilation tube was used for insertion. Myringotomy group: • adenoidectomy, myringotomy and suction of middle ear fluid on both ears. Myringotomy was done with myringotomy knife in the anteroinferior quadrant of tympanic membrane. Interventions used in both groups: • All children received systemic antibiotics, analgesics, anti-inflammatory and decongestant nasal drops for 7 postoperative days.
Outcomes	Final hearing threshold at 12 months (air-bone gap). Tympanic membrane perforation Persistence of OME at 12 months Adverse events
Funding sources	Kerala State board of medical research
Declarations of interest	No competing interests are declared.
Notes	**Research integrity checklist:** No retraction notices or expressions of concern were identified. No prospective trial registration was identified. Baseline characteristics were not excessively similar. Full follow-up was reported. No implausible results were noted. Equal numbers of participants were allocated to each group.

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Quote: “They were randomized into group A and group B as per randomisation table.”
Allocation concealment (selection bias)	Unclear risk	Comment: No information on how/whether the allocation sequence was concealed.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: There was no mention of whether the trial was open or blinded. It is therefore assumed to be open. Outcomes could be influenced by a lack of blinding.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: There was no mention of whether the trial was open or blinded. It is therefore assumed to be open. Outcomes could be influenced by a lack of blinding.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: full follow-up was reported.
Selective reporting (reporting bias)	Unclear risk	Comment: A trial protocol was not available for assessment.
Other bias	Unclear risk	Comment: Potential detection bias as the accuracy and reliability of tympanometry, PTA and otoscopy were not reported.

Tao 2020

Study characteristics	
Methods	2-arm, randomised, parallel-group, open controlled trial with randomisation by child and 12 months of follow-up.

Participants	**Location:** China, single centre **Setting of recruitment and treatment:** ENT Department, Guangzhou Women and Children's Medical Center. **Study dates:** January 2016 to June 2018 **Sample size:** • **Number randomised:** 178 [90 in VT plus adenoidectomy group, 88 in myringotomy plus adenoidectomy group] • **Number completed:** 169 [87 in VT plus adenoidectomy group, 82 in myringotomy plus adenoidectomy group] **Participant (baseline) characteristics:** Age (years): VT plus adenoidectomy mean 7.0 (SD 1.9) years; LM plus adenoidectomy mean 7.2 (SD 2.4) years Gender: VT M 42/87 (48%) F 45/87 (52%); LM 42/82 (51%) F 40/82 (49%) **Inclusion criteria:** • Bilateral otitis media with effusion diagnosed by air-drum otoscopy and confirmed by acoustic impedance examination (Type B); • electronic nasopharyngoscopy- confirmed adenoid hypertrophy blocking more than 1/2 of the posterior nares • middle ear effusion persisting longer than 3 months after conservative treatment, which includes nasal corticosteroids, oral montelukast sodium, oral muco-active agents, and modified Eustachian tube insufflation, plus added antibiotics if complicated by acute sinusitis • average bilateral hearing threshold exceeding 25 dB HL for 500, 1 000, 2 000, and 4 000 Hz • patients aged 4 to 12 years **Exclusion criteria:** • A previous history of nose, ear, or nasopharyngeal surgery • cleft palate or other congenital malformations that may affect the state of the middle ear • congenital or acquired immune deficiency • sensorineural hearing loss or mixed hearing loss.
Interventions	**Ventilation tube:** • myringotomy was performed to suck out the intratympanic fluid, and then a conical short-acting silicon middle ear ventilation tube was placed **Myringotomy:** • myringotomy was performed under the otoendoscope, the intratympanic fluid was sucked out. **Interventions administered to both groups:** • Low temperature plasma radiofrequency ablation of the adenoids was performed, which was assisted by indirect nasopharyngoscopy with entry through the mouth, taking care to avoid damage to the Torus tubarius and the pharyngeal opening of the Eustachian tube
Outcomes	Persistent perforation. Persistence of OME - these data were not used in the review, as data were only reported for one group at 3 months of follow-up, and data from later time points will be affected by the use of different additional treatments in each arm. Adverse events.
Funding sources	Not reported.
Declarations of interest	
Notes	**Research integrity checklist:** No retraction notices or expressions of concern were noted. No prospective trial registration was identified. Baseline characteristics were not excessively similar between the two groups. Plausible loss to follow-up was reported.

	No implausible results were found. Different numbers of participants were allocated to each group.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: “All patients were randomly divided into two groups, namely Group A and B, according to the sequence generated by a computer program when they were admitted to the hospital.”
Allocation concealment (selection bias)	Unclear risk	Comment: Insufficient information to assess.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Comment: There was no report of blinding. Blinding of patients and personnel may not have been feasible for operative interventions. However, lack of blinding could influence outcomes.
Blinding of outcome assessment (detection bias) All outcomes	High risk	Comment: There was no report of blinding. Blinding of patients and personnel may not have been feasible for operative interventions. However, lack of blinding could influence outcome interpretation.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Comment: Low attrition rate.
Selective reporting (reporting bias)	Unclear risk	Comment: No protocol was available to assess.
Other bias	Unclear risk	Comment: Insufficient detail in the report to assess whether an important risk of bias exists.

TARGET 2000

Study characteristics

Methods	3 arm, multi centre parallel group RCT, with randomisation by child and 2 year follow-up. For this review we have included data relevant to the comparison of ventilation tube insertion with watchful waiting. Additional data on adenoidectomy are relevant to a companion review (https://doi.org/10.1002/14651858.CD015252).
Participants	**Location:** UK, 11 sites **Setting of recruitment and treatment:** Otorhinolaryngology Departments **Study dates:** April 1994 to January 1998 **Sample size:** • **Number randomised:** 376 [126 Bilateral VT (VTs), 128 VT with Adenoidectomy (VTs + ad), 122 watchful waiting (WW)] • **Number completed:** 321 [109 Bilateral VT (VTs), 109 VT with Adenoidectomy (VTs + ad), 103 watchful waiting (WW) **Participant (baseline) characteristics:** Age (mean (SD) months): VTs 62.5 (10.2), VTs + ad 64.5 (10.3), WW 62.9 (10.4) Gender: VTs M 60/126 (48%) F 66/126 (52%), VTs + ad M 61/128 (48%) F 67/128 (52%), WW M 62/122 (51%) F 60/122 (49%) Hearing threshold at baseline (at visit 2) (mean (SD) dB): VTs 32.2 (6.0), VTs + ad 31.7 (6.4), WW 33.5 (6.4) AOM episodes (> 6 per year) : VTs 5/126 (4%), VTs + ad 5/127 (4%), WW 8/122 (7%) **Inclusion criteria:** • children aged between 3.25 and 6.75 years • referred primarily for otological or hearing reasons • first visit, with no previous ear or adenoid surgery • bilateral type B + B or B + C2 tympanogram combination • better ear HL > 20 dB HL averaged across 0.5, 1, 2 and 4 kHz and air–bone gap > 10 dB • criteria met on two qualifying visits separated by a 12-week period of watchful waiting. **Exclusion criteria:**

	• children with cranio-facial structural abnormalities, severe systemic disease (e.g. diabetes) and non-OME ear disease (e.g. perforation) • where consultant or parent was unduly concerned over a child's speech/language, behaviour, otalgia or nose/throat problems, the child could be managed outside TARGET. • previous VT/adenoid surgery, outside age limits, not accompanied by parent/guardian, other medical exclusion, significant family language problems, parent refusing to take part in study, child unable/unwilling to do audiometry, administrative problems, family/social reasons and protocol mishaps, particularly early in the trial.
Interventions	**Bilateral VTs:** Bilateral Shepard VTs were inserted (http://www.invotec.net/ventilation_tubes.html) following myringotomy and fluid aspiration **Bilateral VT with adenoidectomy:** Bilateral ventilation tubes were inserted, as above, and adenoidectomy was performed by curettage **Watchful waiting (WW):** Children were not allocated to any surgery. However, over the 2-year follow-up period 57% of participants in this group actually underwent surgery.
Outcomes	Mean final hearing threshold • Air conduction thresholds at 0.5, 1.0, 2.0 and 4.0 kHz in each ear at every visit were summarised as the 4-frequency average binaural hearing thresholds Mean change in hearing from baseline Adverse events: • perforation • haemorrhage • tympanosclerosis • functioning VT
Funding sources	Medical Research Council. Trial Registration Number: ISRCTN35793977.
Declarations of interest	Authors reported "None to declare".
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable for earliest publications (published before 2010). Registration was noted for the most recent publication. Baseline characteristics were not excessively similar between the groups. Plausible loss to follow-up was reported. No implausible results. Numbers allocated to each group are not identical.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Quote: ""For each centre, the first five children were randomised according to a computer-generated random number sequence. Thereafter, the minimisation procedure balanced the treatment allocations across four dichotomous factors: boy, girl; <5.25, >5.25 years old at initial visit; manual, non-manual occupation of head of household and baseline hearing <25 dB HL, >25 dB HL."
Allocation concealment (selection bias)	Low risk	Quote: "Randomisation was performed by telephone call from the nurse/ research assistant to the statistician at the MRC Institute of Hearing Research and allocation immediately communicated to the parent," and "This basis of minimisation was not divulged to centres and may be regarded as completely concealed."
Blinding of participants and personnel (performance bias) All outcomes	High risk	No information provided on blinding of participants and personnel. There is a strong possibility that participants and personnel can identify which treatment a participant received and hence change their behaviour as a result.

Blinding of outcome assessment (detection bias) All outcomes	Low risk	“Audiometry was performed by audiologists, independently of the otolaryngologist and research nurse. Clinic pressures meant that these testers, whilst not blinded in the strictest sense, were not aware of the child’s allocation, nor in a position to be influenced by such information were it present.”
Incomplete outcome data (attrition bias) All outcomes	Unclear risk	Losses to follow-up were 55/376 randomised (14.6%) overall with 19/122 (15.6%) in the medical management group, 17/126 (13.5%) in the VT group and 19/128 (14.8%) in the VT+Ad group. Complete data were available for only 76/122 (62.3%), 85/126 (67.5%) and 92/128 (71.9%) in the medical management, VT and VT+Ad groups respectively. Reasons for losses to follow-up after randomisation were not reported.
Selective reporting (reporting bias)	Unclear risk	The trial entry on ISRCTN registry states that “general health, economic impact, behavioural assessment and quality of life” will be assessed. Data on these are published (no economic data) but no details given of the scales used to assess the outcomes.
Other bias	High risk	The trial registration was retrospectively published, raising the possibility of publication bias. In addition, this was an MRC funded, multi centre trial and yet not all outcomes stated in the trial registration were published.

To 1984

Study characteristics	
Methods	2 arm, RCT, randomisation by ear (split body trial), with and at least 12 months follow-up (mean follow-up of 2 years (range 1 to 5 years).
Participants	**Location:** UK **Setting of recruitment and treatment:** no details given **Study dates:** March 1976 to June 1982 **Sample size:** 54 children • **Number randomised:** [54 ears in intervention, 54 ears in comparison] • **Number completed:** [54 ears in intervention, 54 ears in comparison] **Participant (baseline) characteristics:** Age (mean): 7 years and 6 months (range 47 months to 14 years) Gender: M 29/54 (54%): F 25/54 (46%) Duration of disease: not reported but mean follow-up before operation: 7.2 months Treatment used before trial entry: unspecified "medical measures" **Inclusion criteria:** • children under the age of 14 years • presented with secretory otitis media which failed to respond to "medical measures" • reviewed to confirm the “chronic nature of the condition as shown both clinically and by persistently abnormal audiograms and tympanograms". **Exclusion criteria:** • children with asymmetrical hearing losses, in whom the mean hearing levels on the 2 sides showed a difference of more than 6 dB. • children who had grommets inserted for established complications of the disease, such as retraction pockets and obvious thinning of the drum.
Interventions	**Ventilation tube:** insertion of a Shepherd grommet. 22 in the better ear* (9 right and 13 left), 25 in the worse ear* (11 right and 14 left), 7 in which both ears were equal (2 right and 5 left) (*where these refer to comparisons of audiograms) **Myringotomy:** "most participants" had myringotomy in the contralateral ear. **Use of additional interventions:** All participants received adenoidectomy if adenoids had not previously been removed (n = 9), and were present (n = 1 no adenoids).
Outcomes	**Primary outcome:** hearing level **Secondary outcomes:** Adverse events: perforation, retraction segments, tympanosclerosis
Funding sources	Not reported.

Declarations of interest	No declarations are made.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration not applicable (published before 2010). Baseline characteristics are not relevant (split-body trial) No loss to follow-up was reported. No implausible results. The number randomised to each group was identical as this was a split-body trial.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	“Those who did not respond were submitted to the removal of adenoids (if present) and the insertion of a Shepard grommet in one ear chosen at random.” No information is provided about the process used for randomly selecting an ear.
Allocation concealment (selection bias)	Unclear risk	"Those who did not respond were submitted to the removal of adenoids (if present) and the insertion of a Shepard grommet in one ear chosen at random.” No information is provided about concealment of allocation.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	Unclear risk	“The patients were under the care of 2 consultants working independently and the results were reviewed by an independent observer.” It is unclear if this means that the observer was blinded to group allocation, or was simply a separate assessor.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Follow-up appears to be 100% at 12 months for hearing threshold data. Adverse events are reported at later follow-up times but no information is provided on how many had dropped out. It appears that the number of dropouts after 1 year could have been many: "Twenty-three children have been discharged from follow-up having been well and with normal ears for about a year; some of them have had further surgical treatment on one or both sides. The mean follow-up for this group is 27 months." For adverse event outcomes, the RoB for this domain is high.
Selective reporting (reporting bias)	Unclear risk	No protocol or trial registration was found. The published paper reports all expected outcomes.
Other bias	High risk	“In the other ear, myringotomy was usually performed; those cases in the present trial in which myringotomy was not performed were not considered to introduce a significant variation, as Bennett & Chakraborty showed that myringotomy did not produce a more beneficial effect than adenoidectomy alone.” As the contralateral ear was sometimes treated with myringotomy, and sometimes not, it is unclear whether the study really compared a VT to no treatment, or to myringotomy.

Velepic 2011

Study characteristics	
Methods	Parallel group, single centre RCT with 6 months follow-up. Randomisation by child, analysis by ear. This trial randomised participants to received ventilation tubes and adenoidectomy, or adenoidectomy alone. However, those in the adenoidectomy group were also offered ventilation tube insertion after 3 months, if appropriate. Therefore, we have included this as a comparison of early ventilation tube insertion versus watchful waiting.
Participants	**Location:** Croatia, single centre **Setting of recruitment and treatment**: ENT clinic.

	Study dates: 2004 to 2010 **Sample size:** • **Number randomised:** 161 ears (59 for VT and adenoidectomy, 102 for adenoidectomy alone) • **Number completed:** Not stated, results indicate full follow-up A total of 87 children were included in the study, indicating that most had bilateral disease. **Participant (baseline) characteristics:** Age, years: • VT plus adenoidectomy: mean 5.56 years • Adenoidectomy alone: mean 5.44 years Gender • In total, 37 girls and 50 boys. **Inclusion criteria:** • documented unilateral or bilateral CSOM lasting at least 3 months **Exclusion criteria:** • previous adenoidectomy or tonsillectomy • previous implantation of tympanostomy tubes • craniofacial malformations • congenital ear malformations • chronic otitis media • coagulation disorders. • presence of clinical pathological changes on the structures of the eardrum, including: dangerous attic retractions type III and IV degree, malleus rotation with its drawing closer to, touching, or adhering to the promontorium, first stage of atelectasis of the cavum tympani with retraction pockets of the pars tensa, eardrum adhesion to the incudostapedial joint, or other structures of the medial wall of the cavum.
Interventions	**Ventilation tube plus adenoidectomy:** Operations were performed under GA. Adenoidectomy was performed using Beckmann's adenotome. Myringotomy was performed under the control of operational microscope. It included incision in the posteroinferior quadrant of the eardrum. After the incision, the effusion was aspirated and the tube was inserted. If during the follow-up period CSOM had recurred, the tubes were reinserted. **Adenoidectomy alone:** Participants underwent adenoidectomy. However, If there was no resolution of the effusion after 3 months, myringotomy and implantation of ventilation tube(s) was performed. It is not clear how many participants in this group actually underwent VT tube insertion.
Outcomes	Final hearing threshold • Assessed using the pure tone average air-bone gap across four frequencies. The authors report 'post-operative' measurements. It appears that these were made 'at least 6 months after surgery', but the exact timing is not specified. It is likely, therefore, that at least some participants in the control group had also undergone ventilation tube insertion by this time. Adverse event • persistent perforation • attic retraction • tensa retraction/malleus rotation • scars of the ear drum • myringosclerosis Proportion of children with persistence of OME, identified using "eardrum examination with an operational microscope".
Funding sources	"There was no sponsorship for this study".

Declarations of interest	"Authors report no conflict of interest in the publication of the article. There were no financial and personal relationships with other people or organizations that could inappropriately influence (bias) their work."
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration was not identified (published in 2011). No excessive similarities in baseline characteristics. No loss to follow-up was reported. No implausible results. The number randomised to each group was not identical.

Risk of bias

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	"Children were randomly divided into two groups depending on the treatment method". No details on how the allocation sequence was generated provided. We note a large discrepancy in the number of ears allocated to each group, and this is not explained in the article.
Allocation concealment (selection bias)	Unclear risk	No details on allocation concealment provided.
Blinding of participants and personnel (performance bias) All outcomes	High risk	Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	High risk	No information on blinding of outcome assessors provided for any of the assessments, and the outcomes are not sufficiently objective to discount the possibility of ascertainment bias.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Data missing on one ear (1/161). No information given as to how many children/ears completed the trial.
Selective reporting (reporting bias)	High risk	No protocol or trial registration was found. The published paper reports all expected outcomes, however results are not reported separately per group for adverse events outcomes (although p values have been provided). It is unclear whether outcome data are provided for follow-up at 3 months or 6 months. The time of follow-up would affect interpretation of the outcomes due to the insertion of tympanostomy tubes for all participants in the no tympanostomy tube group who did not have resolution of the effusion after 3 months.
Other bias	High risk	"For 87 children, 37 girls and 50 boys, their parents had signed an informed consent and had regularly come to check-ups. Those children were enrolled in the research." There is the possibility of selection bias as authors chose children who had regularly come to check ups and the outcomes for these children may be different to outcomes for those children who do not regularly attend. A follow-up of six months may be too short to detect a true effect of each intervention.

Yousaf 2016

Study characteristics	
Methods	Parallel group single centre RCT with 6 month follow-up. Randomisation by child.
Participants	**Location:** Pakistan, single centre **Setting of recruitment and treatment:** ENT clinic in Pakistan. **Study dates:** February 2012 to January 2015 **Sample size:** • **Number randomised:** not clear. Apparently 82 participants. • **Number completed:** 82 participants (40 to ventilation tubes, 42 to laser myringotomy)

	Participant (baseline) characteristics: None reported. **Inclusion criteria:** • Diagnosis of unilateral or bilateral OME (diagnostic criteria not described) • Decreased hearing due to persistent middle ear effusion for 6 months or more, “despite three conservative treatments” • Hearing level was more than 30dB • Type B tympanogram • Aged 4-12 years **Exclusion criteria:** • not reported
Interventions	**VT** A myringotomy lancet was used to create an opening in for the insertion of ventilation tubes in the intervention group n=40 children (68 ears) **Laser myringotomy** Performed using an operating microscope. A diode laser of 980nm wavelength with a fibre-optic delivery system was used to perform the myringotomy. The opening was made in the anteroinferior quadrant of the tympanic membrane with a 0.6mm bare diode fibre, projecting 3mm from the hand piece edge. Laser energy was delivered with 5 shots in a circular manner with power of 5 W in 0.5 seconds single-pulse mode. The size of the opening varied from 2 to 2.5mm. n=42 children (68 ears)
Outcomes	Improvement in hearing (definition unclear) Final hearing threshold (for a subset only with persistent effusion) Change in hearing threshold (for a subset only) Adverse events • persistent perforation • persistence of OME • retraction of tympanic membrane • hypertrophic scar • otorrhoea • extrusion of VT
Funding sources	Not reported.
Declarations of interest	No declaration is made.
Notes	**Research Integrity Checklist:** No retraction notices identified. Prospective registration was not identified. Baseline characteristics are not reported. Follow-up was apparently complete. No implausible results. The number of children randomised to each group was not identical (although the number of ears included was identical).

Risk of bias

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Unclear risk	“These patients were randomly allocated to either of the 2 groups.” No information is provided regarding generation of the randomisation sequence. The inclusion of identical numbers of affected ears in each group, despite apparent randomisation at the level of the individual child raises some concerns about the randomisation process.
Allocation concealment	Unclear risk	No details on allocation concealment provided.

(selection bias)		
Blinding of participants and personnel (performance bias) All outcomes	High risk	Surgeons could not be blinded. There is a strong possibility that personnel can identify which treatment a participant received and hence change their behaviour as a result.
Blinding of outcome assessment (detection bias) All outcomes	High risk	There was no report of blinding to treatment allocation for any assessment. The outcomes are not sufficiently objective to discount the possibility of ascertainment bias.
Incomplete outcome data (attrition bias) All outcomes	Low risk	Information on loss to follow-up is not reported, although percentage data for all outcomes indicate no loss to follow-up
Selective reporting (reporting bias)	High risk	No registered protocol was identified, therefore we are unable to compare reported results to pre-specified analysis plan. Hearing was reportedly assessed with pure tone audiogram and tympanogram, but is insufficiently reported, with only the number “improved” in each group, and no clear explanation of what constitutes improvement.
Other bias	High risk	Randomisation seems to have occurred at the level of the individual child. Therefore those with bilateral disease received the same intervention to both ears. However, results are reported at the level of the individual ear. This fails to account for correlation between the ears in the outcome, and may over-estimate the precision of the estimates.

Characteristics of excluded studies [ordered by study ID]

Study	Reason for exclusion
Ah-Tye 2001	ALLOCATION: randomisation not retained
Ardehali 2008	INTERVENTION: treatment with antibiotics, and is relevant for another review in this suite (https://doi.org/10.1002/14651858.CD015254).
Black 1990	PARTICIPANTS: unknown duration of OME
Bozkurt 2004	ALLOCATION: not randomised
Bulman 1984	PARTICIPANTS: wrong patient population. Unknown duration of OME.
Choung 2008	INTERVENTION: treatment with steroids, and is relevant for another review in this suite (https://doi.org/10.1002/14651858.CD015255).
Demant 2017	OTHER: study withdrawn/terminated
El Begermy 2022	PARTICIPANTS: unclear duration of OME.
Englender 1999	ALLOCATION: not randomised
Ferrara 2005	ALLOCATION: not randomised
Gebhart 1981	PARTICIPANTS: wrong patient population (recurrent acute otitis media).
Gibson 1996	ALLOCATION: not randomised
Hammaren-Malmi 2005	PARTICIPANTS: did not have OME of at least 3 months duration
Hao 2019	INTERVENTION: treatment with adenoidectomy, and is relevant for another review in this suite (https://doi.org/10.1002/14651858.CD015252).
Hassmann 2004	ALLOCATION: not randomised
Iino 1989	ALLOCATION: not randomised
Jabeen 2019	INTERVENTION: treatment with adenoidectomy, and is relevant for another review in this suite (https://doi.org/10.1002/14651858.CD015252).
Kremer 1979	ALLOCATION: not randomised
Kujala 2012	PARTICIPANTS: had recurrent acute otitis media, not OME.
Li 2020	COMPARISON: balloon dilatation of the Eustachian tube (inappropriate comparator).
Lildholdt 1983	PARTICIPANTS: unknown duration of OME
Liu 2004	ALLOCATION: not randomised
Mandel 1989	PARTICIPANTS: wrong patient population
Mandel 1992	PARTICIPANTS: wrong patient population
Marchisio 1998	INTERVENTION: treatment with antibiotics, and is relevant for another review in this suite (https://doi.org/10.1002/14651858.CD015254).
Markou 2004	PARTICIPANTS: unknown duration of OME

Study	Reason for exclusion
Maw 1993	INTERVENTION: patients had adenotonsillectomy
Moller 1990	INTERVENTION: treatment with antibiotics, and is relevant for another review in this suite (https://doi.org/10.1002/14651858.CD015254).
MRC Multicentre Otitis Media Study 2004	ALLOCATION: not randomised
MRC Multicentre Otitis Media Study 2008	ALLOCATION: not randomised
NCT00629694	PARTICIPANTS: unknown duration of OME
NCT05545345	INTERVENTION: treatment with adenoidectomy, and is relevant for another review in this suite (https://doi.org/10.1002/14651858.CD015252).
Nguyen 2004	PARTICIPANTS: patients with AOM as well as OME
Paradise 1990	PARTICIPANTS: patients had RAOM
Paradise 1997	ALLOCATION: not randomised
Parlea 2012	ALLOCATION: not randomised
Rohail 2006	PARTICIPANTS: unknown duration of OME
Sanyaolu 2020	ALLOCATION: not randomised
Shishegar 2007	PARTICIPANTS: wrong patient population
Shubich 1996	ALLOCATION: not randomised
Skinner 1988	PARTICIPANTS: wrong patient population
Stenstrom 2005	ALLOCATION: not randomised
Tao 2004	COMPARISONS: wrong intervention
Uvarova 2001	ALLOCATION: not randomised
Xu 2016	INTERVENTION: treatment with adenoidectomy, and is relevant for another review in this suite (https://doi.org/10.1002/14651858.CD015252).
Yousaf 2014	COMPARISONS: comparing two types of myringotomy
Youssef 2013	ALLOCATION: not randomised

Characteristics of studies awaiting classification [ordered by study ID]

Diacova 2016

Methods	—
Participants	—
Interventions	—
Outcomes	—
Notes	Extensive efforts to obtain full text were unsuccessful. The available text is ambiguous in that it defines the design as ‘a prospective observational study’ but then goes on to describe random treatment assignment.

Marshak 1980

Methods	—
Participants	—
Interventions	—
Outcomes	—
Notes	Unable to obtain full-text

Maw 1986

Methods	—
Participants	—
Interventions	—
Outcomes	—
Notes	Unable to obtain full-text

Tawfik 2002

Methods	—
Participants	—
Interventions	—
Outcomes	—
Notes	Unable to obtain full-text

Characteristics of ongoing studies [ordered by study ID]

ACTRN12611001073998

Study name	
Methods	RCT Australia, multi centre 12 month follow-up
Participants	Children with chronic OM
Interventions	Adenoidectomy with VT Adenoidectomy with myringotomy
Outcomes	Trial registration 2011 No data published as of August 2022.
Starting date	
Contact information	
Notes	

NCT02546518

Study name	A Comparison of Surgical and a New Non-Surgical Treatment Methods for Secretory Otitis Media in Children
Methods	Parallel group RCT
Participants	80 children with unilateral or bilateral secretory otitis media of at least 3 months duration, and an intact tympanic membrane.
Interventions	Ventilation tubes compared to Moniri Otovent (autoinflation device).
Outcomes	Change in hearing level measured using age suitable audiogram (1 month, 3 months, 6 months) Change in middle ear pressure using tympanometry (1 month, 3 months, 6 months) Presence of fluid in the middle ear, assessed with otomicroscopy (1 month, 3 months, 6 months) Health economics - number of days of parental leave needed (6 months) Otitis Media Questionnaire-14 (1 month, 3 months, 6 months) Number of healthcare or hospital visits with ear-related issues (6 months)
Starting date	April 2017
Contact information	Mohammed Al-Azzawe: mohammed.al-azzawe@vgregion.se Hasse Ejnell: hasse.ejnell@vgregion.se
Notes	

NCT04584073

Study name	Secretory otitis media in adenoids hypertrophy patients
Methods	Randomised trial, 3 month follow-up.
Participants	**Location:** Egypt **Setting of recruitment and treatment:** ENT department, University hospital **Study dates:** October 2020 to December 2022 (estimated) **Sample size:** • **Estimated enrolment :**150 participants (50 per group)

	Inclusion criteria: • Any case presented with Secretory Otitis Media with adenoids hypertrophy with the following criteria • Age is between 3 to 17 years old • With or without chronic tonsillitis • conductive hearing loss • Recurrent upper respiratory tract infection • Dull tympanic membrane on otoscopy (absent cone of light), decreased mobility of tympanic membrane • Type B tympanogram on tympanometry • OME not responding to medical treatment for three months **Exclusion criteria:** • Patients with the following criteria will be excluded from the study • Previous Myringotomy with or without Tympanostomy Tube application • Previous adenoidectomy or tonsillectomy • Previous ear surgery, cleft palate, Down's syndrome, congenital malformation of the ear and cholesteatoma.
Interventions	1. Adenoidectomy 2. Adenoidectomy and myringotomy 3. Adenoidectomy and myringotomy and tympanostomy tube application
Outcomes	Primary Outcome Measures 1. Tympanogram: 3 months post-surgery 2. Audiogram: 3 months post-surgery
Starting date	October 2020
Contact information	Dr Ahmed Ayman Ahmed Ahmed.20123777@med.au.edu.eg Professor Ahmed Abd El-Hay El-Hussiney alhussiniahmad@aun.edu.eg
Notes	

Appendices

Appendix 1. Search strategies

The search strategies were designed to identify all relevant studies for a suite of reviews on various interventions for otitis media with effusion.

CENTRAL (CRS)	Cochrane ENT Register (CRS)	Medline (Ovid)
1 MESH DESCRIPTOR Otitis Media with Effusion EXPLODE ALL AND CENTRAL:TARGET 2 ("otitis media" adj6 effusion):AB,EH,KW,KY,MC,MH,TI,TO AND CENTRAL:TARGET 3 (OME):TI,TO AND CENTRAL:TARGET 4 (Secretory otitis media):AB,EH,KW,KY,MC,MH,TI,TO AND CENTRAL:TARGET 5 (Serous otitis media):AB,EH,KW,KY,MC,MH,TI,TO AND CENTRAL:TARGET 6 (Middle-ear effusion):AB,EH,KW,KY,MC,MH,TI,TO AND CENTRAL:TARGET	1 MESH DESCRIPTOR Otitis Media EXPLODE ALL AND INREGISTER 2 ("otitis media" OR OME OR "glue ear" OR middle-ear effusion OR middle-ear perfusion):AB,EH,KW,KY,MC,MH,TI,TO AND INREGISTER 3 #1 OR #2 4 (effusion or Recurrent or persistent or serous or secretory or perfusion):AB,EH,KW,KY,MC,MH,TI,TO AND INREGISTER 5 #3 AND #4	1 exp Otitis Media with Effus 2 ("otitis media" adj6 effusion 3 OME.ti. 4 Secretory otitis media.ab,ti 5 Serous otitis media.ab,ti. 6 Middle-ear effusion.ab,ti. 7 Glue ear.ab,ti. 8 middle-ear perfusion.ab,ti. 9 Otitis Media/ 10 otitis media.ti. 11 9 or 10 12 ((effusion or Recurrent or or serous or secretory or per adj3 otitis).ab,ti. 13 11 and 12

7 (glue ear):AB,EH,KW,KY,MC,MH,TI,TO AND CENTRAL:TARGET 8 (middle-ear perfusion):AB,EH,KW,KY,MC,MH,TI,TO AND CENTRAL:TARGET 9 MESH DESCRIPTOR Otitis Media AND CENTRAL:TARGET 10 (otitis media):TI,TO AND CENTRAL:TARGET 11 #9 OR #10 AND CENTRAL:TARGET 12 (((effusion or Recurrent or persistent or serous or secretory or perfusion) adj3 otitis)):AB,EH,KW,KY,MC,MH,TI,TO AND CENTRAL:TARGET 13 #11 AND #12 AND CENTRAL:TARGET 14 #1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #13 AND CENTRAL:TARGET		14 1 or 2 or 3 or 4 or 5 or 6 o 13 15 randomized controlled tria 16 controlled clinical trial.pt. 17 randomized.ab. 18 placebo.ab. 19 drug therapy.fs. 20 randomly.ab. 21 trial.ab. 22 groups.ab. 23 15 or 16 or 17 or 18 or 19 21 or 22 24 exp animals/ not humans 25 23 not 24 26 14 and 25
Embase (Ovid)	**Web of Science (Web of knowledge)**	**Trial registries (CR**
1 exp secretory otitis media/ 2 ("otitis media" adj6 effusion).ab,ti. 3 OME.ti. 4 Secretory otitis media.ab,ti. 5 Serous otitis media.ab,ti. 6 Middle-ear effusion.ab,ti. 7 glue ear.ab,ti. 8 middle-ear perfusion.ab,ti. 9 otitis media/ 10 otitis media.ti. 11 9 or 10 12 ((effusion or Recurrent or persistent or serous or secretory or perfusion) adj3 otitis).ab,ti. 13 11 and 12 14 1 or 2 or 4 or 5 or 6 or 7 or 8 or 13 15 (random* or factorial* or placebo* or assign* or allocat* or crossover*).tw. 16 (control* adj group*).tw. 17 (trial* and (control* or comparative)).tw. 18 ((blind* or mask*) and (single or double or triple or treble)).tw. 19 (treatment adj arm*).tw. 20 (control* adj group*).tw. 21 (phase adj (III or three)).tw. 22 (versus or vs).tw. 23 rct.tw. 24 crossover procedure/ 25 double blind procedure/ 26 single blind procedure/ 27 randomization/ 28 placebo/	11 #10 AND #9 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 10 #8 OR #7 OR #6 OR #5 OR #4 OR #3 OR #2 OR #1 Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 9 TS=(randomised OR randomized OR randomisation OR randomisation OR placebo* OR (random* AND (allocat* OR assign*)) OR (blind* AND (single OR double OR treble OR triple))) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 8 (TI=(otitis media)) AND TS= ((effusion or Recurrent or persistent or serous or secretory or perfusion) NEAR/3 otitis) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 7 TOPIC: ((middle-ear perfusion)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 6 TOPIC: ((glue ear)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 5 TOPIC: ((Middle-ear effusion)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 4 TOPIC: ((Serous otitis media)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 3 TOPIC: ((Secretory otitis media)) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 2 TITLE: (OME)	1 ("otitis media" OR OME OR ear" OR middle-ear effusion middle-ear perfusion):AB,EH,KW,KY,MC AND CENTRAL:TARGET 2 (effusion or Recurrent or p serous or secretory or perfusion):AB,EH,KW,KY,MC AND CENTRAL:TARGET 3 #1 AND #2 4 http*:SO AND CENTRAL:T 5 (NCT0* or ACTRN* or ChiC DRKS* or EUCTR* or eudra IRCT* or ISRCTN* or JapicC JPRN* or NTR0* or NTR1* o NTR3* or NTR4* or NTR5* o NTR7* or NTR8* or NTR9* o or UMIN0*):AU AND CENTRAL:TARGET 6 #4 OR #5 7 #3 AND #6

29 exp clinical trial/ 30 parallel design/ 31 Latin square design/ 32 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 33 exp ANIMAL/ or exp NONHUMAN/ or exp ANIMAL EXPERIMENT/ or exp ANIMAL MODEL/ 34 exp human/ 35 33 not 34 36 32 not 35 37 14 and 36	Indexes=SCI-EXPANDED, CPCI-S Timespan=All years 1 TOPIC: ("otitis media" NEAR/6 effusion) Indexes=SCI-EXPANDED, CPCI-S Timespan=All years	
ClinicalTrials.gov	**ICTRP**	
(EXPAND[Concept] "otitis media" OR EXPAND[Concept] "glue ear" OR middle-ear) AND (effusion OR Recurrent OR persistent OR serous OR secretory OR perfusion) | Interventional Studies	(otitis media AND effusion) OR glue ear OR middle-ear effusion OR middle-ear perfusion	

Appendix 2. Tool for screening eligible studies for scientific integrity/trustworthiness

This screening tool has been developed by Cochrane Pregnancy and Childbirth. It includes a set of predefined criteria to select studies that, based on available information, are deemed to be sufficiently trustworthy to be included in the analysis.

Criteria questions	**Assessment**		**Comments and concerns**
	High risk	**Low risk**	
Research governance			
Are there any retraction notices or expressions of concern listed on the Retraction Watch Database relating to this study?	Yes	No	
Was the study prospectively registered (for those studies published after 2010) If not, was there a plausible reason?	No	Yes	
When requested, did the trial authors provide/share the protocol and/or ethics approval letter?	No	Yes	
Did the trial authors engage in communication with the Cochrane Review authors within the agreed timelines?	No	Yes	
Did the trial authors provide IPD data upon request? If not, was there a plausible reason?	No	Yes	
Baseline characteristics			
Is the study free from characteristics of the study participants that appear too similar? (e.g. distribution of the mean (SD) excessively narrow or excessively wide, as noted by Carlisle 2017)	No	Yes	
Feasibility			
Is the study free from characteristics that could be implausible? (e.g. large numbers of women with a rare condition (such as severe cholestasis in pregnancy) recruited within 12 months)	No	Yes	
In cases with (close to) zero losses to follow-up, is there a plausible explanation?	No	Yes	
Results			
Is the study free from results that could be implausible? (e.g. massive risk reduction for main outcomes with small sample size)?	No	Yes	
Do the numbers randomised to each group suggest that adequate randomisation methods were used (e.g. is the study free from issues such as unexpectedly even numbers of women ‘randomised’ including a mismatch between the numbers and the methods, if the authors say ‘no blocking was	No	Yes	

used' but still end up with equal numbers, or if the authors say they used 'blocks of 4' but the final numbers differ by 6)?			
For abstracts only:			
Have the study authors confirmed in writing that the data to be included in the review have come from the final analysis and will not change?	No	Yes	

Appendix 3. Additional detail on adverse effects

Comparison 1: Ventilation tubes versus no treatment

VT vs no treatment

Rach 1991 found that in the short term (< 3 months) 9/44 (20.5%) VT were *in situ* and in the medium term (6 months) 18/44 (40.9%) of the tubes had extruded in the VT only group (assessed by otoscopy).

Maw 1983 reports that some VTs were reinserted, but no data are presented for the number of extrusions/reinsertions.

Dempster 1993 reported that at 12 months tympanosclerosis had occurred in 28 (39%) of ears in the VT group but in none of the ears without VT. In addition, at 12 months, 6 (8.3%) ears in the VT and 7 (9.7%) ears in the no treatment group showed signs of perforation/retraction. At the 12 months follow-up visit, 31% of VT were still functioning.

Comparison 2: Ventilation tubes versus watchful waiting (WW)

In the TARGET 2000 trial, of 635 ears that had a VT inserted, eight had a perforation recorded at least 6 months after surgery. However, of the 4 who attended later appointments, all had healed. Of ears receiving a VT, either with and without adenoidectomy, 128/635 (20%) showed tympanosclerosis while none were reported in the watchful waiting group. For ears receiving VT, in the short term, 259/327 ears (79%) were functioning while 68/327 (21%) were either non-functioning or extruded, in the medium term (12 months) 57/316 ears (55%) were functioning while 259/316 (18%) were either non-functioning or extruded and in the long term (24 months) 9/300 ears (3%) were functioning while 291/300 (97%) were either non-functioning or extruded. Data are presented only for ears when the otoscopy and tympanometry results agree. One child (1/165 (0.6%) who underwent an adenoidectomy had to return to theatre for postoperative haemorrhage (*Note: the total number exceeds the number allocated to adenoidectomy because of cross-overs from other groups*).

Maw 1999 did not report adverse events.

Paradise 2007 assessed assessed a number of adverse events after long term follow-up. The results were as follows:

- Tympanosclerosis
    - RR 0.91 for those undergoing early ventilation tube insertion (95% CI 0.33 to 2.55; 1 study; 391 participants, but data adjusted to account for non-independence of within-individual measurement; Analysis 2.16; very low certainty evidence).
- Fibrosis
    - RR 0.61 for those undergoing early ventilation tube insertion (95% CI 0.10 to 3.60; 1 study; 391 participants, but data adjusted to account for non-independence of within-individual measurement; Analysis 2.17; very low certainty evidence).
- Segmental atrophy
    - RR 2.83 for those undergoing early ventilation tube insertion (95% CI 1.81 to 4.43; 1 study; 391 participants, but data adjusted to account for non-independence of within-individual measurement; Analysis 2.18; very low certainty evidence).

- Retraction pocket with other abnormality
    - RR 0.91 for those undergoing early ventilation tube insertion (95% CI 0.06 to 14.50; 1 study; 391 participants, but data adjusted to account for non-independence of within-individual measurement; Analysis 2.19 very low certainty evidence).

Rach 1991 did not report adverse events after long-term follow-up (relevant for this comparison).

Rovers 2000 presented data on the proportion of children with parental reports of otorrhoea in the short term (3 months), with 42.9% in the VT group and 14.3% in the WW group. In the medium term (12 months) 37.6% in the VT group reported otorrhoea while 16.5% did in the WW group. Rovers 2000 also reported the number of children with a specific number of episodes of otorrhoea at 12 months. In the VT group 16/93 (17%) of children reported number of episodes of otorrhoea, 28 (30%) reported one episode, 26 (28%) reported two episodes and 23 (25%) reported more than three episodes. In the WW group, 58 (62%) reported no episodes of otorrhoea at 12 months, 23 (24%) reported one episode, 8 (9%) reported two episodes and 5 (5%) reported three episodes. In terms of cumulative proportion of children with one or more episodes of otorrhoea at 12 months, 83% in the VT group (95% CI 75 to 91%) and 38% (28 to 48%) in the WW group (P=0.001). At three months 92% of VT were *in situ*, and 30% at 12 months.

Velepic 2011 presented data for a number of adverse events but data were presented for all participants rather than for each group. In terms of attic retractions 74/161 (46%) ears presented as mild retractions (type I and II according to Sudhoff and Tos), while in 5/161 (3.1%) ears retractions were severe (type III and IV). A total of 82/161 (51%) ears showed no attic retraction. Velepic 2011 reported that when the two groups were compared, ears in the adenoidectomy only group more frequently reported normal ears in term of attic retraction compared to ears receiving adenoidectomy and VT (chi-square=4.592; ss=1; p=0.032). Tensa retractions/malleus rotation was observed in 36/161 ears(22.4%). There was no statistically significant difference in the incidence between the two groups (chi-square=0.263; ss=1; p=0.608). Scars of the ear drum were observed in 46/161 ears (28.6%) and were found significantly more frequently in the group receiving VT (chi-square=28.107; ss=1; p<0.001). Myringosclerosis was observed in 42/161 ears (26.1%) but there was no significant difference in the incidence observed between the two groups (chi-square=0.171; ss=1; p=0.680). Data on persistent perforation are shown in Analysis 2.9.

Comparison 3: Ventilation tubes versus myringotomy

All adverse events reported by Bernard 1991 are included in Table 3 and Table 4. Comparative data were available for myringosclerosis, with a risk ratio of 4.60 for those who received ventilation tubes (95% CI 1.64 to 12.91; 1 study; 125 participants; Analysis 3.3; very low-certainty evidence).

Comparison 4: Ventilation tubes versus myringotomy

In the D'Eredita 2006 trial, participants were asked to report “any complications noted during the post-operative period” in a questionnaire. D'Eredita 2006 reported that 59 of 60 questionnaires (98.3%) were returned. Given that there were 30 children participating in the trial, it is not clear whether participants were asked to complete one questionnaire on two occasions for each child or one questionnaire for each ear on one occasion. It is therefore not clear whether the adverse events reported relate to children or ears. Parents reported six episodes of otorrhoea: two in the laser myringotomy group at two months post surgery, and four in the VT group at 30 days and 3 months post surgery. The otorrhoea responded to topical antibiotic containing drops.

Gates 1989 reported necrosis of the long process of the incus in one child who received a VT and the child underwent a myringostapediopexy. It is not clear to which treatment group the child was randomised. A tube fell into the middle ear in three instances and became trapped when the tympanic membrane healed. In such cases, repeat myringotomy was performed, the tube removed and a new one inserted. The time point of

assessment was not stated but assumed to be two years. Gates 1989 reported the number (proportion) of children with the number of episodes of otorrhoea (see Analysis 4.14).

Koopman 2004 reported that 1/208 (0.5%) children in the LM group complained of severe otalgia during the first 2 days post laser myringotomy. There were no signs of inflammation, and the condition was treated with oral analgesics. Otorrhoea occurred more frequently in the VT ear than in the laser myringotomy ear (p=0.002) but the number of events and denominators were not reported.

Popova 2010 reported episodes of otorrhoea per child at the medium term (12 months). For children receiving adenoidectomy and VT 25/ 42 (60%) reported no episodes of otorrhoea, 10/ 42 (24%) reported one episode, 5/ 42 (12%) reported two episodes, 1/ 42 (2%) reported three episodes and 1/ 42 (2%) reported four or more episodes. In the children receiving adenoidectomy and myringotomy, all children 36/ 36 (100%) reported no episodes of otorrhoea. Of the 42 children receiving VT, 7 (17%) experienced a blockage.

Ruckley 1988 found no evidence of tympanosclerosis in any ear receiving either treatment. In the short term (3 months) 2/36 (5.5%) of ears receiving VT were blocked. In the very short term (2 weeks), one child complained of mild otalgia in the ear receiving thermal myringotomy ear. Persistent perforation Analysis 4.5

Sujatha 2015 reported adverse events by ear. In the right ear, in the group receiving myringotomy alone, 22(88%) showed retracted TM at 3 months, and at one year 7(28%) were retracted and 1(4%) showed tympanosclerotic patch. In those receiving VT at one year 14(56%) were retracted, 2(8%) showed tympanosclerotic patch and 3(12%) TM showed perforation in the anterior quadrant. This is significant by Fishers exact test (p<0. 01). (Fig. 3)

In the left ear, in the group receiving myringotomy alone. after one year, 6(24%) showed retracted TM whereas those receiving VT showed retraction in 12 (48%) cases, tympanosclerotic patch in 1(4%) and perforation in 3(12%). All perforations were in the anterior quadrant. This comparison between groups showed significant difference by Fishers exact test (p<0. 05).

In the right ear;. all VT was in situ at third month visit and all but one expelled at the end of 6months. In the left ear, VT was present in all patients in the 3rd month follow-up and it was expelled in all except one at the 6th month visit. In one case VT got blocked at 3rd month and it was removed under local anaesthesia.

Tao 2020 reported that at 2 weeks follow-up, of those receiving myringotomy, 5 ears/4 patients showed tympanic effusion while in those receiving VT non-purulent effusions could be seen in the ear canals in 8 ears/7 patients and the re-examination after 1 week showed that all the ears were dry. A re-examination 6 months after operation showed that in those receiving myringotomy 3 ears/2 patients received tympanostomy again and at 12 months, 2 ears/2 patients received tympanostomy again after the failure of conservative treatment.

To 1984 reported that 9/54 (17%) receiving a VT experienced tympanosclerosis while 1/54 (2%) ears receiving a myringotomy experienced tympanosclerosis. The timing of the follow-up was not reported. In terms of retraction segments, 0/54 ears receiving VT and 1/54 receiving a myringotomy experienced retraction segments assessed at 9 months, while 2/54 (4%) ears receiving VT and 1/54 receiving a myringotomy experienced retraction segments assessed in the long term (24 months). In terms of persistent perforation, 1 ear receiving VT experienced this between 9 and 21 months and 0 ears receiving myringotomy). Analysis 4.7

Yousaf 2016. In terms of post surgical haemorrhage those receiving LM reported 0 cases but 9 (13%) in the VT group reported this. Yousaf 2016 reported that for ears receiving VT 6/68 (13%) had extruded in the very short term (30 days) while 53/68 (78%) had extruded in the medium term (6 months).

- Retraction of the tympanic membrane: RR 2.33 for those receiving ventilation tubes as compared to laser myringotomy (95% CI 0.64 to 8.46; 1 study; 90

participants; Analysis 4.17; very low-certainty evidence). Sensitivity analysis to account for correlation between ears of the same individual made little difference to the overall effect estimates (Analysis 7.15; Analysis 7.16).

- Hypertrophic scar of the tympanic membrane: OR 7.50 for those receiving ventilation tubes as compared to laser myringotomy (95% CI 0.46 to 121.15; 1 study; 90 participants; Analysis 4.18; very low-certainty evidence)
- Otorrhoea: RR 3.00 for those receiving ventilation tubes as compared to laser myringotomy (95% CI 0.32 to 27.76; 1 study; 90 participants; Analysis 4.19; very low-certainty evidence). Sensitivity analysis to account for correlation between ears of the same individual made little difference to the overall effect estimates (Analysis 7.17; Analysis 7.18).

References

References to studies included in this review

Bernard 1991 {published data only}

Bernard PA, Stenstrom RJ, Feldman W, Durieux-Smith A. Randomized, controlled trial comparing long-term sulfonamide therapy to ventilation tubes for otitis media with effusion. Pediatrics 1991;88(2):215-22. [PMID: 1861917]

Bernard PA, Stenstrom RJ, Feldman W, Durieux-Smith A. Randomized, controlled trial comparing long-term sulfonamide therapy to ventilation tubes for otitis media with effusion. Pediatrics 1991 Aug;88(2):215-22. [CENTRAL: CN-00077175] 1012090 [PMID: 1861917]

Feldman W, Bernard P, Smith A, Stenstrom R. Sulfonamide prophylaxis vs ventilation tubes in hearing-loss due to recurrent otitis-media with effusion (Rome) - preliminary-results of a randomized controlled trial. Clinical and Investigative Medicine-medecine Clinique et Experimentale 1987;(4):A33. [CENTRAL: CN-02494379] 6720430

Feldman W, Bernard P, Smith A, Stenstrom R. Sulfonamide prophylaxis vs ventilation tubes in hearing-loss due to recurrent otitis-media with effusion - preliminary-results of a randomized controlled trial. American Journal of Diseases of Children 1987;(4):389-9. [CENTRAL: CN-02494378] 6720424

D'Eredita 2006 {published data only}

D'Eredità R, Shah UK. Contact diode laser myringotomy for medium-duration middle ear ventilation in children. International journal of pediatric otorhinolaryngology 2006;70:1077-80. [DOI: 10.1016/j.ijporl.2005.11.003]

Dempster 1993 {published data only}

* Dempster JH, Browning GG, Gatehouse SG. A randomized study of the surgical management of children with persistent otitis media with effusion associated with a hearing impairment. Journal of laryngology and otology 1993;107(4):284-9. [DOI: 10.1017/s0022215100122844]

Dempster JH, Browning GG, Gatehouse SG. A randomized study of the surgical management of children with persistent otitis media with effusion associated with a hearing impairment. Journal of Laryngology and Otology 1993 Apr;107(4):284-9. [CENTRAL: CN-00094188] 1028834 [PMID: 8320510]

Elkholy 2021 {published data only}

Elkholy TA, El-gaber A, Mohammed F, Al-Agamy DMM, Elhady M. Impact of adenoidectomy alone or with ventilation tube for treatment of secretory otitis media in children. Al-azhar International Medical Journal 2021;2(8):53-8. [CENTRAL: CN-02518323] 22462018

Gates 1989 {published data only}

Gates GA, Avery CA, Cooper JC, Prihoda TJ, Cooper JCJ. Chronic secretory otitis media: effects of surgical management. Annals of Otology, Rhinology & Laryngology. Supplement 1989 Jan;138:2-32. [CENTRAL: CN-00057471] 992514 [PMID: 2492178]

Gates GA, Avery CA, Prihoda TJ, Cooper JC. Effectiveness of adenoidectomy and tympanostomy tubes in the treatment of chronic otitis media with effusion. New England Journal of Medicine 1987 Dec 3;317(23):1444-51. [CENTRAL: CN-00051110] 986159 [PMID: 3683478]

Gates GA, Avery CA, Prihoda TJ. Effect of adenoidectomy upon children with chronic otitis media with effusion. Laryngoscope 1988 Jan;98(1):58-63. [CENTRAL: CN-00051547] 986595 [PMID: 3336263]

Gates GA, Wachtendorf C, Hearne EM, Holt GR. Treatment of chronic otitis media with effusion: results of tympanostomy tubes. American Journal of Otolaryngology 1985 May-Jun;6(3):249-53. [CENTRAL: CN-00038728] 973784 [PMID: 4040338]

Koopman 2004 {published data only}

[DOI: 10.1097/00005537-200405000-00010]

Koopman JP, Blom H, Hoeve H, Mulder P. Laser myringotomy versus ventilation tubes in children with chronic otitis media with effusion: a randomized trial for efficacy and outcome prediction. In: 8th International Symposium on Recent Advances in Otitis Media. 2003:136. [CENTRAL: CN-00449322] 1302469

Koopman JP, Reuchlin AG, Kummer EE, Boumans LJ, Rijntjes E, Hoeve LJ, et al. Laser myringotomy versus ventilation tubes in children with otitis media with effusion: a randomized trial. Laryngoscope 2004 May;114(5):844-9. [CENTRAL: CN-00469076] 1318610 [PMID: 15126741]

Maw 1983 {published data only}

Maw AR, Herod F. Otoscopic, impedance, and audiometric findings in glue ear treated by adenoidectomy and tonsillectomy. A prospective randomised study. Lancet (london, England) 1986 Jun 21;1(8495):1399-402. [CENTRAL: CN-00043077] 978132 [PMID: 2872514]

Maw AR, Parker A. Surgery of the tonsils and adenoids in relation to secretory otitis media in children. Acta Oto-laryngologica. Supplementum 1988;454:202-7. [CENTRAL: CN-00057952] 992995 [PMID: 3223250]

Maw AR. Age and adenoid size in relation to adenoidectomy in otitis media with effusion. American Journal of Otolaryngology 1985 May-Jun;6(3):245-8. [CENTRAL: CN-00038727] 973783 [PMID: 4040337]

Maw AR. Chronic otitis media with effusion (glue ear) and adenotonsillectomy: prospective randomised controlled study. British Medical Journal (clinical Research Ed.) 1983 Nov 26;287(6405):1586-8. [CENTRAL: CN-00032849] 967912 [PMID: 6416513]

Maw AR. Chronic otitis media with effusion and adeno-tonsillectomy--a prospective randomized controlled study. International Journal of Pediatric Otorhinolaryngology 1983 Dec;6(3):239-46. [CENTRAL: CN-00034535] 969597 [PMID: 6373644]

Maw AR. Factors affecting adenoidectomy for otitis media with effusion (glue ear). Journal of the Royal Society of Medicine 1985 Dec;78(12):1014-8. [CENTRAL: CN-00040567] 975623 [PMID: 4067973]

Maw AR. Otitis media with effusion (glue ear). A prospective randomized study 1 year following adenoidectomy and tonsillectomy. Clinical Otolaryngology and Allied Sciences 1984;9:130. [CENTRAL: CN-00262086] 1161207

Maw AR. The long term effect of adenoidectomy on established otitis media with effusion in children. Auris, Nasus, Larynx 1985;12(Suppl 1):S234-6. [CENTRAL: CN-00043248] 978303 [EMBASE: 16731316] [PMID: 3835918]

Parker AJ. No peak-peak tympanometric conversion in relation to radiographic palatal airway size in children with otitis media with effusion (OME). Clin-Otolaryngol

1989;14:172. [CENTRAL: CN-00262390] 1161497

Maw 1999 {published data only}

Hall AJ, Maw AR, Steer CD. Developmental outcomes in early compared with delayed surgery for glue ear up to age 7 years: a randomised controlled trial. Clinical Otolaryngology 2009 Feb;34(1):12-20. [CENTRAL: CN-00699827] 1502988 [PMID: 19260880]

Maw R, Wilks J, Harvey I, Peters TJ, Golding J, Wiks J. Early surgery compared with watchful waiting for glue ear and effect on language development in preschool children: a randomised trial. Lancet (london, England) 1999 Mar 20;353(9157):960-3. [CENTRAL: CN-00166722] 1097611 [PMID: 10459904]

Maw R, Wilks J, Harvey I, Peters TJ, Golding J, Wiks J. Early surgery compared with watchful waiting for glue ear and effect on language development in preschool children: a randomised trial. Lancet 1999;353(9157):960-3.

Maw R, Wilks J, Harvey I, Peters TJ, Golding J. Randomised controlled trial of early surgery versus watchful waiting for glue ear: the effect of langauge development in pre-school children. Journal of Laryngology 1999;113(23 Suppl):40. [CENTRAL: CN-00292554] 1181639

Wilks J, Maw R, Peters TJ, Harvey I, Golding J. Randomised controlled trial of early surgery versus watchful waiting for glue ear: the effect on behavioural problems in pre-school children. Clinical Otolaryngology and Allied Sciences 2000 Jun;25(3):209-14. [CENTRAL: CN-00330781] 1206154 [PMID: 10944051]

Paradise 2007 {published data only}

[CTG: NCT00365092]

Johnston LC, Feldman HM, Paradise JL, Bernard BS, Colborn DK, Casselbrant ML, et al. Tympanic membrane abnormalities and hearing levels at the ages of 5 and 6 years in relation to persistent otitis media and tympanostomy tube insertion in the first 3 years of life: a prospective study incorporating a randomized clinical trial. Pediatrics 2004 Jul;114(1):e58-67. [CENTRAL: CN-00490067] 1336081 [PMID: 15231974]

NCT00365092. Middle Ear Disease Before Age 3, Treatment With Ear Tubes, and Literacy and Attentional Abilities at Ages 9 to 11 [Early Otitis and Literacy and Attention at 9 to 11 Years]. Https://clinicaltrials.gov/show/NCT00365092 (first received 2006 Aug 15). [CENTRAL: CN-02013311] 12551781

Paradise J, Dollaghan C, Campbell T, Feldman H. Early vs delayed tube placement for persistent middle-ear effusion in the first three years of life: effects on cognition, language, speech and behavior at ages 4 and 6 years. In: 8th International Symposium on Recent Advances in Otitis Media . Fort Lauderdale, FL, USA, 3-7 June, 2003. 2003:219. [ABSTRACT NO.: B57] [CENTRAL: CN-00449362] 1302502

Paradise Jack L, Campbell Thomas F, Dollaghan Christine A, Feldman Heidi M, Bernard Beverly S, Colborn DK, et al. Developmental outcomes after early or delayed insertion of tympanostomy tubes. New England Journal of Medicine 2005 Aug 11;353(6):576-86. [CENTRAL: CN-00523784] 1366143 [EMBASE: 2005369617] [PMID: 16093466]

Paradise JL, Colborn DK, Bernard BS, Hakos LM, Guerra NJ, et al. Comparison of the early results of tympanostomy tube placement and of nonsurgical management in infants with persistent otitis media. In: Recent Advances in Otitis Media . 6th International Symposium on Recent Advances in Otitis Media, Jun 04-08, 1995, Ft Lauderdale, Florida. 1996:201. [CENTRAL: CN-00634260] 1449625

Paradise JL, Dollaghan CA, Campbell TF, Feldman HM, Bernard BS, Colborn DK, et al. Otitis media and tympanostomy tube insertion during the first three years of life: developmental outcomes at the age of four years. Pediatrics 2003 Aug;112(2):265-77. [CENTRAL: CN-00439773] 1294118 [PMID: 12897272]

Paradise JL, Feldman HM, Campbell TF, Dollaghan CA, Colborn DK, Bernard BS, et al. Early versus delayed insertion of tympanostomy tubes for persistent otitis media: developmental outcomes at the age of three years in relation to

prerandomization illness patterns and hearing levels. Pediatric Infectious Disease Journal 2003 Apr;22(4):309-14. [CENTRAL: CN-00431844] 1287293 [PMID: 12690269]

Paradise JL, Feldman HM, Campbell TF, Dollaghan CA, Colborn DK, Bernard BS, et al. Early vs late tube placement for persistent middle-ear effusion (MEE) in the first 3 years of life: effects on language, speech sound production, and cognition at age 3 years. Pediatric Research 2000;47(4):216A. [CENTRAL: CN-00383898] 1249198

Paradise JL, Feldman HM, Campbell TF, Dollaghan CA, Colborn DK, Bernard BS, et al. Effect of early or delayed insertion of tympanostomy tubes for persistent otitis media on developmental outcomes at the age of three years. New England Journal of Medicine 2001 Apr 19;344(16):1179-87. [CENTRAL: CN-00327417] 1202797 [EMBASE: 32319492] [PMID: 11309632]

Paradise JL, Feldman HM, Campbell TF, Dollaghan CA, Rockette HE, Pitcairn DL, et al. Tympanostomy tubes and developmental outcomes at 9 to 11 years of age. New England Journal of Medicine 2007 Jan 18;356(3):248-61. [CENTRAL: CN-00575011] 1404219 [EMBASE: 2007-01205-003] [PMID: 17229952]

Popova 2010 {published data only}

Popova D, Varbanova S, Popov TM. Comparison between myringotomy and tympanostomy tubes in combination with adenoidectomy in 3-7-year-old children with otitis media with effusion. International Journal of Pediatric Otorhinolaryngology 2010 Jul;74(7):777-80. [CENTRAL: CN-00767591] 1565308 [PMID: 20399511]

Rach 1991 {published data only}

* Rach GH, Zielhuis GA, van Baarle PW, van den Broek P. The effect of treatment with ventilating tubes on language development in preschool children with otitis media with effusion. Clinical Otolaryngology and Allied Sciences 1991 Apr;16(2):128-32. [CENTRAL: CN-00076714] 1011629 [PMID: 2070526]

Schilder AGM, van Manen JG, Zielhuis GA, Grievink EH. Long-term results of a randomized controlled trial of surgery for otitis media with effusion. In: Sade J, editors(s). Infections in childhood Ear, Nose and Throat Aspects: proceedings of the 3rd International Conferenceof the European Working Group for Pediatric Otorhinolargyngology, Jerusalem, Israel: Nov 7-12, 1993. 1994.

Zielhuis GA, Rach GH, van den Broek P. Screening for otitis media with effusion in preschool children. Lancet 1989;1(8633):311-4.

Rovers 2000 {published data only}

Hartman M, Rovers MM, Ingels K, Zielhuis GA, Severens JL, van der Wilt GJ. Economic evaluation of ventilation tubes in otitis media with effusion. Archives of Otolaryngology--head & Neck Surgery 2001 Dec;127(12):1471-6. [CENTRAL: CN-00369148] 1235952 [EMBASE: 2001440522] [PMID: 11735817]

Rovers M, Ingels K, van der Wilt GJ, van den Broek P, Zielhuis G. The effects of ventilation tubes: a randomized clinical trial. In: Recent Advances in Otitis Media, Proceedings. 2001:411-4. [CENTRAL: CN-02494418] 18710324

Rovers M, Zielhuis G, Hartman M, Ingels K, Straatman H, Van Der Wilt G. The effect of early screening and treatment with ventilation tubes in infants with persistent otitis media with effusion. In: ISTAHC Annual Meeting. Vol. 16. 2000:063. [CENTRAL: CN-00452876] 1305273

Rovers MM, Krabbe PF, Straatman H, Ingels K, van der Wilt GJ, Zielhuis GA. Randomised controlled trial of the effect of ventilation tubes (grommets) on quality of life at age 1-2 years. Archives of Disease in Childhood 2001 Jan;84(1):45-9. [CENTRAL: CN-00327716] 1203095 [PMID: 11124783]

Rovers MM, Straatman H, Ingels K, van der Wilt GJ, van den Broek P, Zielhuis GA. Generalizability of trial results based on randomized versus nonrandomized allocation of OME infants to ventilation tubes or watchful waiting. Journal of

Clinical Epidemiology 2001 Aug;54(8):789-94. [CENTRAL: CN-00349653] 1220304 [PMID: 11470387]
Rovers MM, Straatman H, Ingels K, van der Wilt GJ, van den Broek P, Zielhuis GA. The effect of short-term ventilation tubes versus watchful waiting on hearing in young children with persistent otitis media with effusion: a randomized trial. Ear and Hearing 2001 Jun;22(3):191-9. [CENTRAL: CN-00362609] 1231427 [PMID: 11409855]
Rovers MM, Straatman H, Ingels K, van der Wilt GJ, van den Broek P, Zielhuis GA. The effect of ventilation tubes on language development in infants with otitis media with effusion: a randomized trial. Pediatrics 2000 Sep;106(3):E42. [CENTRAL: CN-00299394] 1185683 [PMID: 10969126]
Rovers MM, Zielhuis K, Ingels H, Straatman G. The effect of early screening and treatment with ventilation tubes in infants with persistent otitis media with effusion. Clinical Otolaryngology and Allied Sciences 2003;28(2):158. [CENTRAL: CN-00452877] 1305274

Ruckley 1988 {published data only}

Ruckley RW, Blair RL. Thermal myringotomy (an alternative to grommet insertion in childhood secretory otitis media?). Journal of Laryngology and Otology 1988;102(2):125-8. [CENTRAL: CN-00452878] 1305275

Sujatha 2015 {published data only}

Sujatha S, Venugopalan PG. Evaluation of effectiveness of myringotomy and myringotomy with ventilation tube insertion as treatment of otitismedia with effusion in children. Journal of Evolution of Medical and Dental Sciences 2015;4(29):5003-5009.

Tao 2020 {published data only}

Tao J, Luo R, Chen Y, Hou C, Hao Q. Myringotomy or tympanostomy tube insertion, comparison of surgical treatment of adenoid hypertrophy and otitis media with effusion in children. Lin Chung Er Bi Yan Hou Tou Jing Wai Ke Za Zhi 2020;34(3):207-210. [DOI: 10.13201/j.issn.2096-7993.2020.03.005] [PMID: 32791583]

TARGET 2000 {published data only}

Browning G. TARGET trial of alternative regimes for glue ear treatment: analysis 2. In: 8th International Congress of Paediatric Otorhinolaryngology (ESPO) . Oxford, UK, 11-14 September, 2002. 2002:60. [CENTRAL: CN-00431582] 1287083
Browning GG. Patient-centred oucomes of a surgical RCT of otitis media with effusion at one year. Laryngo-Rhino-Otologie 2000;79(Suppl 1):S35-36. [CENTRAL: CN-00526378] 1368375
Browning GG. Two-year outcome of ventilation tubes in a randomized controlled trial of persistent childhood otitis media with effusion. Clinical Otolaryngology and Allied Sciences 2001;26(4):342-4. [CENTRAL: CN-02494417] 1304904
Haggard MP, Gannon MM, Birkin JA, Bennett KE, Nicholls EE, Spencer H, et al. Adjuvant adenoidectomy in persistent bilateral otitis media with effusion: hearing and revision surgery outcomes through 2 years in the TARGET randomised trial. Clinical Otolaryngology 2012 Apr;37(2):107-16. [CENTRAL: CN-00896356] 1668857 [EMBASE: 364670938] [PMID: 22443163]
ISRCTN35793977. Trial of alternative regimens in glue ear treatment - effectiveness of surgery for otitis media with effusion in 3.5-7 year olds using multiple developmental and economic measures combined with classical clinical measures. Http://www.controlled-trials.com/ISRCTN35793977 (first received 2000). [CENTRAL: CN-00725114] 1526959
MRC Multicentre Otitis Media Study Group. Surgery for persistent otitis media with effusion: generalizability of results from the UK trial (TARGET). Trial of Alternative Regimens in Glue Ear Treatment. Clinical Otolaryngology and Allied Sciences

2001 Oct;26(5):417-24. [CENTRAL: CN-00374837] 1240427 [EMBASE: 32995195] [PMID: 11678951]

MRC Multicentre Otitis Media Study Group. The role of ventilation tube status in the hearing levels in children managed for bilateral persistent otitis media with effusion. Clinical Otolaryngology and Allied Sciences 2003 Apr;28(2):146-53. [CENTRAL: CN-00431744] 1287213 [EMBASE: 2003174240] [PMID: 12680834]

To 1984 {published data only}

To SS, Pahor AL, Robin PE. A prospective study of the use of unilateral grommets in chronic bilateral secretory otitis media. Clinical Otolaryngology and Allied Sciences 1984;9:130-1. [CENTRAL: CN-00262088] 1161209

To SS, Pahor AL, Robin PE. A prospective trial of unilateral grommets for bilateral secretory otitis media in children. Clinical Otolaryngology and Allied Sciences 1984 Apr;9(2):115-7. [CENTRAL: CN-00035295] 970357 [PMID: 6380827]

Velepic 2011 {published data only}

Velepic M, Starcevic R, Bonifacic M, Ticac R, Kujundzic M, Udovic DS, et al. The clinical status of the eardrum: an inclusion criterion for the treatment of chronic secretory otitis media in children. International Journal of Pediatric Otorhinolaryngology 2011 May;75(5):686-90. [CENTRAL: CN-00784332] 1581204 [EMBASE: 51315541] [PMID: 21397957]

Yousaf 2016 {published data only}

Yousaf M, Malik SA, Haroon T. Laser myringotomy versus ventilation tubes in otitis media with effusion. Journal of Ayub Medical College, Abbottabad 2016 Oct-Dec;28(4):773-5. [CENTRAL: CN-01380598] 6212560 [EMBASE: 619635268] [PMID: 28586587]

References to studies excluded from this review

Ah-Tye 2001 {published data only}

Ah-Tye C, Paradise JL, Colborn DK. Otorrhea in young children after tympanostomy-tube placement for persistent middle-ear effusion: prevalence, incidence, and duration. Pediatrics 2001 Jun;107(6):1251-8. [CENTRAL: CN-00348427] 1219078 [PMID: 11389239]

Ardehali 2008 {published data only}

Ardehali MM, Seraj JM, Asiabar MK, Adibi H. The possible role of gastroesophageal reflux disease in children suffering from chronic otitis media with effusion. Acta Medica Iranica 2008;46(1):33-7. [CENTRAL: CN-00708224] 1511305 [EMBASE: 351792703]

Black 1990 {published data only}

Black NA, Sanderson CF, Freeland AP, Vessey MP. A randomised controlled trial of surgery for glue ear. BMJ (Clinical Research Ed.) 1990 Jun 16;300(6739):1551-6. [CENTRAL: CN-00069039] 1004056 [PMID: 2196954]

Bozkurt 2004 {published data only}

Bozkurt MK, Calguner M. The efficacy of CO2 laser myringotomy in serous otitis media. Kulak Burun Bogaz Ihtisas Dergisi : KBB [Journal of Ear, Nose, and Throat] 2004;12(3-4):55-9. [CENTRAL: CN-00523128] 1365490 [PMID: 16010101]

Bulman 1984 {published data only}

Bulman CH, Brook SJ, Berry MG. A prospective randomized trial of adenoidectomy vs grommet insertion in the treatment of glue ear. Clinical Otolaryngology and Allied Sciences 1984 Apr;9(2):67-75. [CENTRAL: CN-00174952] 1103203 [EMBASE: 1984176580] [PMID: 6380828]

Choung 2008 {published data only}

Choung YH, Shin YR, Choi SJ, Park K, Park HY, Lee JB, et al. Management for the children with otitis media with effusion in the tertiary hospital. Clinical and Experimental Otorhinolaryngology 2008;1(4):201-5. [CENTRAL: CN-00671567] 1478067

Demant 2017 {published data only}

[CTG: NCT02490332]

Demant MN, Jensen RG, Jakobsen JC, Gluud C, Homoe P. The effects of ventilation tubes versus no ventilation tubes for recurrent acute otitis media or chronic otitis media with effusion in 9 to 36 month old Greenlandic children, the SIUTIT trial: study protocol for a randomized controlled trial. Trials 2017 Jan 19;18(1):30. [CENTRAL: CN-01304999] 4901463 [EMBASE: 614107422] [PMID: 28103950]

NCT02490332. The Effects of Ventilation Tubes - The SIUTIT Trial [The Effects of Ventilation Tubes Versus no Ventilation Tubes for Recurrent Acute Otitis Media or Chronic Otitis Media With Effusion in 9 to 36 Month Old Greenlandic Children - a Randomised Clinical Trial]. Https://clinicaltrials.gov/show/NCT02490332 (first received 2015 Jul 3). [CENTRAL: CN-02043426] 12581896

El Begermy 2022 {published data only}

El Begermy MA, El Begermy MM, Kassamy H. The effect of endoscopic peritubal adenoidectomy vs myringotomy with ventilation tubes insertion in management of otitis media with effusion in children. Egyptian Journal of Ear, Nose, Throat and Allied Sciences 2022;23(23):1-10. [CENTRAL: CN-02518322] 22462011

Englender 1999 {published data only}

Englender M, Somech E, Harell M. Laser myringotomy (L-myringotomy) and ventilating tubes: a preliminary comparative study. Lasers in Medical Science 1999;14(1):62-6. [CENTRAL: CN-00431622] 1287110

Ferrara 2005 {published data only}

Ferrara S, Sammartano D, Ferrara P. Long-term management of recurrent otitis media with effusion in children. In: XVIII IFOS World Congress; 2005 Jun 25-30, Rome (Italy). 2005. [CENTRAL: CN-00526409] 1368400

Gebhart 1981 {published data only}

Gebhart DE. Tympanostomy tubes in the otitis media prone child. Laryngoscope 1981 Jun;91(6):849-66. [CENTRAL: CN-00025323] 960394 [EMBASE: 1981143267] [PMID: 7017311]

Gibson 1996 {published data only}

Gibson PG, Stuart JE, Wlodarczyk J, Olson LG, Hensley MJ. Nasal inflammation and chronic ear disease in Australian Aboriginal children. Journal of Paediatrics and Child Health 1996 Apr;32(2):143-7. [CENTRAL: CN-00131589] 1066176 [PMID: 8860389]

Hammaren-Malmi 2005 {published data only}

Hammaren-Malmi S, Saxen H, Tarkkanen J, Mattila PS. Adenoidectomy does not significantly reduce the incidence of otitis media in conjunction with the insertion of

tympanostomy tubes in children who are younger than 4 years: a randomized trial. Pediatrics 2005 Jul;116(1):185-9. [CENTRAL: CN-00528797] 1370430 [PMID: 15995051]

Hao 2019 {published data only}

Hao J, Chen M, Liu B, Yang Y, Liu W, Ma N, et al. Compare two surgical interventions for otitis media with effusion in young children. European Archives of Oto-rhino-laryngology 2019 Aug;276(8):2125-31. [CENTRAL: CN-01956237] 11330475 [EMBASE: 627898353] [PMID: 31127413]

Hao J, Chen M, Liu B, Yang Y, Liu W, Ma N, et al. Correction to: compare two surgical interventions for otitis media with effusion in young children [Correction to: Compare two surgical interventions for otitis media with effusion in young children (European Archives of Oto-Rhino-Laryngology, (2019), 276, 8, (2125-2131), 10.1007/s00405-019-05421-9)]. European Archives of Oto-rhino-laryngology 2019 Aug;276(8):2133-4. [CENTRAL: CN-02080898] 12863908 [EMBASE: 628234545] [PMID: 31214828]

Hassmann 2004 {published data only}

Hassmann E, Skotnicka B, Baczek M, Piszcz M. Laser myringotomy in otitis media with effusion: long-term follow-up. European Archives of Oto-rhino-laryngology 2004;261(6):316-20. [CENTRAL: CN-00497408] 1342900

Iino 1989 {published data only}

Iino Y, Ishitoya J, Ikeda M, Ito Y, Usami M, Kawashiro N, et al. Factors on delayed recovery of otitis media with effusion in children--clinical and bacteriological study. Nihon Jibiinkoka Gakkai Kaiho 1989 Aug;92(8):1183-91. [CENTRAL: CN-00063873] 998912 [PMID: 2685215]

Jabeen 2019 {published data only}

Jabeen F, Chaudhry S, Ahmed Z, Khalil N, Amin B, Rizvi SS. Comparison of rate of recurrence of otitis media with effusion in children treated by myringotomy and ventilating tube insertion with patients treated by additional adenoidectomy. Rawal Medical Journal 2019;44(3):513-6. [CENTRAL: CN-02007003] 12422789 [EMBASE: 2002854226]

Kremer 1979 {published data only}

Kremer M, Podoshin L, Fradis M. Treatment of serous otitis media with tympanic ventilation tubes. Ear, Nose, & Throat Journal 1979 May;58(5):203-9. 12616379 [PMID: 456299]

Kujala 2012 {published data only}

Kujala T, Alho OP, Kristo A, Uhari M, Renko M, Pokka T, et al. Quality of life after surgery for recurrent otitis media in a randomized controlled trial. Pediatric Infectious Disease Journal 2014 Jul;33(7):715-9. [CENTRAL: CN-00995945] 1726577 [EMBASE: 2014406374] [PMID: 24445832]

Li 2020 {published data only}

Li H, Zhou Y, Huang J, Zhang L, Lou Z. Treatment of Secretory Otitis Media with Balloon Dilation and Tympanostomy Tube Insertion: a Randomized Controlled Trial. Basic & Clinical Pharmacology & Toxicology 2020;127(Suppl 1):292-3. [CENTRAL: CN-02230226] 16641841 [EMBASE: 633830686]

Lildholdt 1983 {published data only}

Lildholdt T. Ventilation tubes in secretory otitis media. A randomized, controlled study of the course, the complications, and the sequelae of ventilation tubes. Acta Oto-

laryngologica. Supplementum 1983;398:1-28. [CENTRAL: CN-00033048] 968111 [EMBASE: 1984015453] [PMID: 6316711]

Liu 2004 {published data only}

Liu L, Sun YG, Ma L, Zhao W, Wu R. Effect of ventilation tube insertion on otitis media with effusion in cleft palate children. Zhonghua Er Bi Yan Hou Ke za Zhi 2004;39(4):216-8. [CENTRAL: CN-00549233] 1379301 [EMBASE: 39714817]

Mandel 1989 {published data only}

Mandel EM, Rockette HE, Bluestone CD, Paradise JL, Nozza RJ. Myringotomy with and without tympanostomy tubes for chronic otitis media with effusion. Archives of Otolaryngology--head & Neck Surgery 1989 Oct;115(10):1217-24. [CENTRAL: CN-00062654] 997695 [PMID: 2789777]

Mandel 1992 {published data only}

Mandel EM, Rockette HE, Bluestone CD, Paradise JL, Nozza RJ. Efficacy of myringotomy with and without tympanostomy tubes for chronic otitis media with effusion. Pediatric Infectious Disease Journal 1992 Apr;11(4):270-7. [CENTRAL: CN-00338034] 1212771 [PMID: 1565550]

Marchisio 1998 {published data only}

Marchisio P, Principi N, Passali D, Salpietro DC, Boschi G, Chetri G, et al. Epidemiology and treatment of otitis media with effusion in children in the first year of primary school. Acta Oto-laryngologica 1998 Jul;118(4):557-62. [CENTRAL: CN-00154484] 1085418 [EMBASE: 1998254893] [PMID: 9726683]

Markou 2004 {published data only}

Markou K, Vlachtsis K, Kyriafinis G, Petridis D, Daniilidis I. Long-term results of adenoidectomy and ventilation tube insertion in the surgical treatment of chronic otitis media with effusion. International Journal of Pediatric Otorhinolaryngology 2004;68(5):684. [CENTRAL: CN-00477460] 1325509

Maw 1993 {published data only}

Maw R, Bawden R, Maw AR. Spontaneous resolution of severe chronic glue ear in children and the effect of adenoidectomy, tonsillectomy, and insertion of ventilation tubes (grommets). BMJ (Clinical Research Ed.) 1993 Mar 20;306(6880):756-60. [CENTRAL: CN-00093062] 1027712 [PMID: 8490338]

Moller 1990 {published data only}

Moller P, Dingsor G. Otitis media with effusion: can erythromycin reduce the need for ventilating tubes? Journal of Laryngology and Otology 1990 Mar;104(3):200-2. [CENTRAL: CN-00067688] 1002714 [PMID: 2341774]

MRC Multicentre Otitis Media Study 2004 {published data only}

MRC Multicentre Otitis Media Study Group. Speech reception in noise: an indicator of benefit from otitis media with effusion surgery. Clinical Otolaryngology and Allied Sciences 2004 Oct;29(5):497-504. [CENTRAL: CN-00497404] 1342896 [EMBASE: 39287082] [PMID: 15373863]

MRC Multicentre Otitis Media Study 2008 {published data only}

MRC Multicentre Otitis Media Study Group. An extension of the Jerger classification of tympanograms for ventilation tube patency--specification and evaluation of equivalent ear-canal volume criteria. Ear and Hearing 2008 Dec;29(6):894-906. [CENTRAL: CN-00682310] 1487454 [EMBASE: 550156802] [PMID: 18685496]

NCT00629694 {published data only}

NCT00629694. Adenoidectomy, Myringotomy and Tubes' Insertion vs Adenoidectomy and Myringotomy Alone in Children With Otitis Media With Effusion and Adenoid Hypertrophy. Https://clinicaltrials.gov/show/NCT00629694 (first received 2008 March 06). [CENTRAL: CN-00726812] 1528205

NCT05545345 {published data only}

NCT05545345. Adjuvant adenoidectomy for the treatment of chronic OME in children [Adjuvant adenoidectomy plus tympanostomy tube placement for the treatment of chronic OME in children]. Https://clinicaltrials.gov/show/NCT05545345 (first received 2022). [CENTRAL: CN-02462316] 21608888

Nguyen 2004 {published data only}

Nguyen LH, Manoukian JJ, Yoskovitch A, Al-Sebeih KH, Al Sebeih KH. Adenoidectomy: selection criteria for surgical cases of otitis media. Laryngoscope 2004 May;114(5):863-6. [CENTRAL: CN-00469074] 1318608 [PMID: 15126745]

Paradise 1990 {published data only}

Paradise JL, Bluestone CD, Rogers KD, Taylor FH, Colborn DK, Bachman RZ, et al. Efficacy of adenoidectomy for recurrent otitis media in children previously treated with tympanostomy-tube placement. Results of parallel randomized and nonrandomized trials. JAMA 1990 Apr 18;263(15):2066-73. [CENTRAL: CN-00066665] 1001697 [PMID: 2181158]

Paradise 1997 {published data only}

Paradise J, Campbell T, Dollaghan C, Feldman H, Bernard B, Colborn K, et al. Receptive vocabulary, cognition, and parent-rated behavior at age 3 years in relation to otitis media in the first 3 years of life. In: Abstract Book of the Association of Health Service Research. Vol. 14. 1997:350-1. [CENTRAL: CN-00452820] 1305231

Parlea 2012 {published data only}

Parlea E, Georgescu M, Calarasu R. Tympanometry as a predictor factor in the evolution of otitis media with effusion. Journal of Medicine and Life 2012 Dec 15;5(4):452-4. [CENTRAL: CN-00850240] 1626099 [EMBASE: 369159817] [PMID: 23346249]

Rohail 2006 {published data only}

Rohail A, Gill ZI, Butt MR. A comparison of medical treatment versus surgical treatment for the management of Otitis media with effusion. Annals of King Edward Medical College 2006;12(1):64-7. [CENTRAL: CN-00597368] 1421949

Sanyaolu 2020 {published data only}

Sanyaolu LN, Cannings-John R, Butler CC, Francis NA. The effect of ventilation tube insertion on quality of life in children with persistent otitis media with effusion. Clinical Otolaryngology 2020 Mar;45(2):239-47. [CENTRAL: CN-02217143] 12616424 [EMBASE: 2004053181] [PMID: 31869494]

Shishegar 2007 {published data only}

Shishegar M, Hoghoghi H. Comparison of adenoidectomy and myringotomy with and without tube placement in the short term hearing status of children with otitis media with effusion: a preliminary report. Iranian Journal of Medical Sciences 2007;32(3):169-72. [CENTRAL: CN-00708699] 1511740 [EMBASE: 351270400]

Shubich 1996 {published data only}

Shubich I. Otitis media with effusion and allergy control in children: a prospective study. In: Sixth International Symposium on Otitis Media; 1996, Fort Lauderdale (FL). 1996:173-4. [CENTRAL: CN-00452904] 1305298

Skinner 1988 {published data only}

Skinner DW, Lesser TH, Richards SH. A 15-year follow-up of a controlled trial of the use of grommets in glue ear. Otolaryngology and Allied Sciences 1989;14:169.

Skinner DW, Lesser THJ, Richards SH. A 15 year follow-up of a controlled trial of the use of grommets in glue ear. Clinical Otolaryngology and Allied Sciences 1988;13(5):341-6. 12618597 [EMBASE: 18269160]

Stenstrom 2005 {published data only}

Stenstrom R, Pless IB, Bernard P. Hearing thresholds and tympanic membrane sequelae in children managed medically or surgically for otitis media with effusion. Archives of Pediatrics & Adolescent Medicine 2005 Dec;159(12):1151-6. [CENTRAL: CN-00532329] 1373954 [PMID: 16330739]

Tao 2004 {published data only}

TAO L-hua, Ding-hua HE, You-di HUANG. Effects of Adenoidectomy in Treatment of Secretory Otitis Media in Children. Journal of Clinical Pediatric Surgery 2004. [CENTRAL: CN-00849652] 1625619

Uvarova 2001 {published data only}

Uvarova N. Comparison of the results of treatment of otitis media with effusion after myringotomy with and without ventiliation tube. In: 4th Extraordinary International Symposium on Recent Advances in Otitis Media; 2001 Apr 16-20, Sendai (Japan). 2001:165. [ABSTRACT NO.: 122] [CENTRAL: CN-00362633] 1231448

Xu 2016 {published data only}

Xu WM, Ye YH. Effect of tympanostomy tube insertion with adenoidectomy for children with recurrent otitis media with effusion. Lin Chuang Er Bi Yan Hou Tou Jing Wai Ke za Zhi [Journal of Clinical Otorhinolaryngology, Head, and Neck Surgery] 2016;30(23):1873-5. [CENTRAL: CN-01614396] 8828012 [EMBASE: 622451696] [PMID: 29798016]

Yousaf 2014 {published data only}

Yousaf M, Malik SA, Zada B. Laser and incisional myringotomy in otitis media with effusion-a comparative study. Journal of Ayub Medical College, Abbottabad 2014 Oct-Dec;26(4):441-3. [CENTRAL: CN-01043373] 1773966 [PMID: 25672160]

Youssef 2013 {published data only}

Youssef TF, Ahmed MR. Laser-assisted myringotomy versus conventional myringotomy with ventilation tube insertion in treatment of otitis media with effusion: Long-term follow-up. Interventional Medicine and Applied Science 2013;5(1):16-20. [CENTRAL: CN-01043520] 1774110 [PMID: 24265883]

References to studies awaiting assessment

Diacova 2016 {published data only}

Diacova S, McDonald TJ, Ababii I. Clinical, functional, and surgical findings in chronic bilateral otitis media with effusion in childhood. Ear, Nose, & Throat Journal 2016 Aug;95(8):E31-7. [CENTRAL: CN-01339417] 5406197 [EMBASE: 614412861] [PMID: 27551851]

Marshak 1980 {published data only}

Marshak G, Neriah ZB. Adenoidectomy versus tympanostomy in chronic secretory otitis media. Annals of Otology, Rhinology, and Laryngology. Supplement 1980;89(3 Part 2):316-8. 3446829

Maw 1986 {published data only}

Maw AR. Adenoidectomy and adenotonsillectomy for otitis media with effusion in children: a prospective randomized controlled study. 1986. [CENTRAL: CN-00849155] 1625223

Tawfik 2002 {published data only}

Tawfik S, Belal A, Sorour W. A comparative study of the different treatment modalities of otitis media with effusion in children. In: 8th International Congress of Paediatric Otorhinolaryngology (ESPO) . Oxford, UK, 11-14 September, 2002. 2002:151. [ABSTRACT NUMBER: P2.26] [CENTRAL: CN-00508402] 1352190

References to ongoing studies

ACTRN12611001073998 {unpublished data only}

ACTRN12611001073998. Surgery for otitis media in Indigenous Australian children [A 12 month, multi-centred, randomized trial to compare the outcomes of two surgical and one medical intervention on chronic Otitis Media in Indigenous children living in remote communities of Australia. Medicine V surgery sub-study]. Http://www.anzctr.org.au/ACTRN12611001073998.aspx (first received 2011 October 17). [CENTRAL: CN-01012938] 1743555

NCT02546518 {published data only}

NCT02546518. A Comparison of Surgical and a New Non-Surgical Treatment Methods for Secretory Otitis Media in Children- Hearing and Socioeconomic Aspects. Https://clinicaltrials.gov/show/nct02546518 (first received 2015). [CENTRAL: CN-01102072] 1832581

NCT04584073 {published data only}

NCT04584073. Secretory Otitis Media in Adenoids Hypertrophy Patients [Comparison Between the Fate of Secretory Otitis Media in Patients With Adenoids Hypertrophy Undergoing Adenoidectomy Alone or With Myringotomy or With Myringotomy and Tympanostomy Tube Application]. Https://clinicaltrials.gov/show/NCT04584073 (first received 2020 Oct 12). [CENTRAL: CN-02181953] 14787741

Additional references

Abidin 1995

Abidin RR. Parenting Stress Index: Professional Manual. 3rd edition. Odessa, FL: Psychological Assessment Resources, 1995.

Achenbach 2011

Achenbach TM. Child Behavior Checklist. In: Kreutzer JS, DeLuca J, Caplan B, editors(s). Encyclopedia of Clinical Neuropsychology. New York, NY: Springer, 2011. [DOI: 10.1007/978-0-387-79948-3_1529]

Bayley 2006

Bayley N. Bayley Scales of Infant and Toddler Development. 3rd edition. San Antonio, TX: Harcourt Assessment, Inc, 2006.

Berkman 2013

Berkman ND, Wallace IF, Steiner MJ, Harrison M, Greenblatt AM, Lohr KN, et al. Otitis Media with Effusion: Comparative Effectiveness of Treatments. Vol. Report No.: 13-EHC091-EF. Rockville (MD): Agency for Healthcare Research and Quality (US), 2013. [PMID: 23762917]

Browning 2010

Browning GG, Rovers MM, Williamson I, Lous J, Burton MJ. Grommets (ventilation tubes) for hearing loss associated with otitis media with effusion in children. Cochrane Database of Systematic Reviews 2010, Issue 10. Art. No: CD001801. [DOI: 10.1002/14651858.CD001801.pub3]

Bruce 2015

Bruce I, Harman N, Williamson P, Tierney S, Callery P, Mohiuddin S, et al. The management of Otitis Media with Effusion in children with cleft palate (mOMEnt): a feasibility study and economic evaluation. Health Technology Assessment 2015;19(68):1-374. [DOI: 10.3310/hta19680]

Cheong 2012

Cheong KH, Hussain SS. Management of recurrent acute otitis media in children: systematic review of the effect of different interventions on otitis media recurrence, recurrence frequency and total recurrence time. Journal of Laryngology & Otology 2012;126(9):874-85.

Cochrane ENT 2020

Cochrane ENT. Otitis media with effusion: a project to prioritise Cochrane systematic reviews. https://ent.cochrane.org/otitis-media-effusion-ome-glue-ear 2020.

de Beer 2004

de Beer B, Schilder AGM, Zielhuis GA, Ingels K, Graamans K. Hearing loss in young adults who had ventilation tube insertion in childhood. Annals of Otology, Rhinology, and Laryngology 2004;113(6):438-44. [DOI: 10.1177/000348940411300604]

de Beer 2005

de Beer BA, Schilder AGM, Zielhuis GA, Graamans K. Natural course of tympanic membrane pathology related to otitis media and ventilation tubes between ages 8 and 18 years. Otology & Neurotology 2005;26(5):1016-21. [DOI: 10.1097/01.mao.0000185058.89586.ed]

Dunn 2007

Dunn LM, Dunn DM. Peabody Picture Vocabulary Test (PPVT™-4). 4th edition. Pearson Education, 2007.

Fekkes 2000

Fekkes M, Theunissen NC, Brugman E, Veen S, Verrips EGH, Koopman HM, et al. Development and psychometric evaluation of the TAPQOL: a health-related quality of life instrument for 1–5-year-old children. Quality of Life Research 2000;9:961-72. [DOI: 10.1023/a:1008981603178]

Flynn 2009

Flynn T, Möller C, Jönsson R, Lohmander A . The high prevalence of otitis media with effusion in children with cleft lip and palate as compared to children without clefts.

International Journal of Pediatric Otorhinolaryngology 2009;73:1441-6. [DOI: 10.1016/j.ijporl.2009.07.015]

Galbraith 2022

Galbraith K, Mulvaney CA, MacKeith S, Marom T, Daniel M, Venekamp RP, et al. Autoinflation for otitis media with effusion (OME) in children. Cochrane Database of Systematic Reviews 2022, Issue 4. Art. No: CD015253. [DOI: 10.1002/14651858.CD015253]

Goodman 1997

Goodman R. The Strengths and Difficulties Questionnaire: a research note. Journal of Child Psychology and Psychiatry 1997;38:581-6. [DOI: 10.1111/j.1469-7610.1997.tb01545.x]

Gresham 1990

Gresham FM, Elliott SN. Social Skills Rating System. Circle Pines, MN: American Guidance Service, 1990.

Griffiths 1996

Griffiths R. The Griffiths mental development scales from birth to two years, manual, the 1996 revision. Henley: Association for Research in Infant and Child Development, Test Agency, 1996.

Haggard 2003

Haggard MP, Smith SC, Nicholls EE. Quality of life and child behaviour. In: Rosenfeld RM, Bluestone CD, editors(s). Evidence-based Otitis Media. 2nd edition. Hamilton, Ontario: BC Decker Inc, 2003:401-29. [https://researchonline.lshtm.ac.uk/id/eprint/15108]

Handbook 2011

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.

Hedrick 1984

Hedrick DL, Prather EM, Tobin AR. Sequenced Inventory of Communication Development. Seattle, WA: University of Washington Press, 1984.

Higgins 2021

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.2 (updated February 2021). Cochrane, 2021. Available from training.cochrane.org/handbook.

Jellinek 1988

Jellinek MS, Murphy JM, Robinson J, Feins A, Lamb S, Fenton T. Pediatric Symptom Checklist: screening school-age children for psychosocial dysfunction. Journal of Pediatrics 1988;112(2):201-9. [DOI: 10.1016/s0022-3476(88)80056-8]

Kreiner-Møller 2012

Kreiner-Møller E, Chawes BL, Caye-Thomasen P, Bønnelykke K, Bisgaard H. Allergic rhinitis is associated with otitis media with effusion: a birth cohort study. Clinical

and Experimental Allergy 2012;42(11):1615-20. [DOI: 10.1111/j.1365-2222.2012.04038.x]

Laina 2006

Laina V, Pothier DD. Should we aspirate middle-ear effusions prior to insertion of ventilation tubes? Journal of Laryngology and Otology 2006;120:818-21. [DOI: 10.1017/S0022215106002118]

Landgraf 1994

Landgraf JM. The Infant/Toddler Child Health Questionnaire: conceptual framework, logic content, and preliminary psychometric results. Boston: Health Act, 1994.

Landgraf 1996

Landgraf JL, Abetz L, Ware JE. The CHQ User's Manual. Boston: The Health Institute, New England Medical Center, 1996.

Lefebvre 2020

Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf M-I, et al. Technical Supplement to Chapter 4: Searching for and selecting studies. Higgins JPT, Thomas J, Chandler J, Cumpston MS, Li T, Page MJ, et al, editors(s). Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (updated September 2020). Cochrane, 2020. Available from training.cochrane.org/handbook.

Liu 2020

Liu PZ, Ismail-Koch H, Stephenson K, Donne AJ, Fergie N, Derry J, et al. A core outcome set for research on the management of otitis media with effusion in otherwise-healthy children. International Journal of Pediatric Otorhinolaryngology 2020 ;134:Article 110029. [DOI: 10.1016/j.ijporl.2020.110029]

MacKeith 2022a

MacKeith S, Mulvaney CA, Galbraith K, Marom T, Daniel M, Venekamp RP, et al. Adenoidectomy for otitis media with effusion (OME) in children. Cochrane Database of Systematic Reviews 2022, Issue 4. Art. No: CD015252. [DOI: 10.1002/14651858.CD015252]

Maris 2014

Maris M, Wojciechowski M, Van de Heyning P, Boudewyns A. A cross-sectional analysis of otitis media with effusion in children with Down syndrome. European Journal of Pediatrics 2014;173:1319-25. [DOI: 10.1007/s00431-014-2323-5]

Marseglia 2008

Marseglia GL, Pagella F, Caimmi D, Caimmi S, Castellazzi AM, Poddighe D, et al. Increased risk of otitis media with effusion in allergic children presenting with adenoiditis. Otolaryngology – Head and Neck Surgery 2008;138(5):572-5. [DOI: 10.1016/j.otohns.2008.01.020]

Marshall 2018

Marshall J, Noel-Storr AH, Kuiper J, Thomas J, Wallace BC. Machine learning for identifying randomized controlled trials: an evaluation and practitioner's guide. Research Synthesis Methods 2018;9(4):602-14. [DOI: 10.1002/jrsm.1287]

McCarthy 1972

McCarthy D. Manual for the McCarthy Scales of Children's Abilities. New York: Psychological Corp, 1972.

McDonald 2017

McDonald S, Noel-Storr AH, Thomas J. Harnessing the efficiencies of machine learning and Cochrane Crowd to identify randomised trials for individual Cochrane reviews. In: Global Evidence Summit; 2017 Sep 13-17; Cape Town, South Africa. 2017.

Mulvaney 2022a

Mulvaney CA, Galbraith K, MacKeith S, Marom T, Daniel M, Venekamp RP, et al. Antibiotics for otitis media with effusion (OME) in children. Cochrane Database of Systematic Reviews 2022, Issue 4. Art. No: CD015254. [DOI: 10.1002/14651858.CD015254]

Mulvaney 2022b

Mulvaney CA, Galbraith K, MacKeith S, Marom T, Daniel M, Venekamp RP, et al. Topical and oral steroids for otitis media with effusion (OME) in children. Cochrane Database of Systematic Reviews 2022, Issue 4. Art. No: CD015255. [DOI: 10.1002/14651858.CD015255]

NICE 2008

NICE. Otitis media with effusion in under 12s: Otitis media with effusion in under 12s: surgery (CG60). Available at: https://www.nice.org.uk/guidance/cg60 2008.

NICE CKS 2021

NICE. Otitis media with effusion (Clinical Knowledge Summary). Available at: https://cks.nice.org.uk/topics/otitis-media-with-effusion/ 2021.

Noel-Storr 2018

Noel-Storr AH. Cochrane Crowd: new ways of working together to produce health evidence. In: Evidence Live; 2018 Jun 18-20; Oxford, UK. 2018.

Rabin 2001

Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Annals of Medicine 2001;33(5):337-43. [DOI: 10.3109/07853890109002087]

RevMan 2020 [Computer program]

Review Manager (RevMan). Version 5.4. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2020.

Reynell 1985

Reynell JH. Reynell Development Language Scales Manual. 2nd edition. Windsor, UK: NFER-NELSON, 1985.

Rosenfeld 1997

Rosenfeld RM, Goldsmith AJ, Tetlus L, Balzano A. Quality of life for children with otitis media. Archives of Otolaryngology - Head & Neck surgery 1997;123:1049-54. [DOI: 10.1001/archotol.1997.01900100019002]

Rosenfeld 2000

Rosenfeld RM, Bhaya MH, Bower CM, Brookhouser PE, Casselbrant ML, Chan KH, et al. Impact of tympanostomy tubes on child quality of life. Archives of

Otolaryngology - Head & Neck Surgery 2000;126:585-92.

Rosenfeld 2003

Rosenfeld RM, Kay D. Natural history of untreated otitis media. Laryngoscope 2003;113:1645-57. [DOI: 10.1097/00005537-200310000-00004]

Rosenfeld 2016

Rosenfeld RM, Shin JJ, Schwartz SR, Coggins R, Gagnon L, Hackell JM, et al. Clinical practice guideline: otitis media with effusion (update). Otolaryngology - Head & Neck Surgery 2016;154:S1-S41. [DOI: 10.1177/0194599815623467]

Schilder 2016

Schilder AGM, Chonmaitree T, Cripps AW, Rosenfeld RM, Casselbrant ML, Haggard MP, et al. Otitis media. Nature Reviews Disease Primers 2016;2(1):16063. [DOI: 10.1038/nrdp.2016.63]

Schlichting 2007

Schlichting JEPT, Lutje Spelberg HC. Lexilijst Begrip: An Instrument to Investigate Language Comprehension in Children aged 15-25 Months in the Context of Early Identification. Amsterdam: Pearson Assessment & Information BV, 2007.

Schlichting 2010

Schlichting JEPT, Lutje Spelberg HC. Schlichting Test for Language Comprehension; Instruction Manual. Woooden: Bohn Stafleu van Loghum, 2010.

Schmalbach 2021

Schmalbach Cecelia E, Brereton Jean, Bowman Cathlin, Denneny III James C. American Academy of Otolaryngology–Head and Neck Surgery/Foundation Regent Registry: Purpose, Properties, and Priorities. Otolaryngology–Head and Neck Surgery 2021;164(5):964-971.

Steele 2017

Steele D, Adam GP, Di M, Halladay C, Pan I, Coppersmith N, et al. Tympanostomy tubes in children with otitis media. Comparative Effectiveness Review No. 185 (Prepared by the Brown Evidence-based Practice Center under Contract No. 290-2015-00002-I.). AHRQ Publication No. 17-EHC003-EF. Rockville, MD: Agency for Healthcare Research and Quality, 2017. [DOI: 10.23970/AHRQEPCCER185]

Thomas 2017

Thomas J, Noel-Storr AH, Marshall I, Wallace B, McDonald S, Mavergames C, et al, Living Systematic Review Network. Living systematic reviews 2: combining human and machine effort. Journal of Clinical Epidemiology 2017;91:31-7. [DOI: 10.1016/j.jclinepi.2017.08.011]

TNO 1997

TNO—Prevention and Health/LUMC. TAIQOL—Questionnaire for parents of children aged 1—5 years. Leiden, The Netherlands: Leiden University Medical Center, 1997.

van den Aardweg 2010

van den Aardweg MTA, Schilder AGM, Herkert E, Boonacker CWB, Rovers MM. Adenoidectomy for otitis media in children. Cochrane Database of Systematic

Reviews 2010, Issue 1. Art. No: CD007810. [DOI: 10.1002/14651858.CD007810.pub2]

Vanneste 2019

Vanneste P, Page C. Otitis media with effusion in children: pathophysiology, diagnosis, and treatment. A review. Journal of Otolaryngology - Head & Neck Surgery 2019;14(2):33-9. [DOI: 10.1016/j.joto.2019.01.005]

Venekamp 2018

Venekamp RP, Mick P, Schilder AGM, Nunez DA. Grommets (ventilation tubes) for recurrent acute otitis media in children. Cochrane Database of Systematic Reviews 2018, Issue 5. Art. No: CD012017. [DOI: 10.1002/14651858.CD012017.pub2]

Verrips 1998

Verrips GH, Vogels AGC, Verloove-Vanhorick SP, Fekkes M, Koopman HM, Kamphuis RP, et al. Health-related quality of life measure for children - the TACQOL. Journal of Applied Therapeutics 1998;1(4):357-60.

Wallace 2014

Wallace IF, Berkman ND, Lohr KN, Harrison MF, Kimple AJ, Steiner MJ. Surgical treatments for otitis media with effusion: a systematic review. Pediatrics 2014;133(2):296-311. [DOI: 10.1542/peds.2013-3228]

Wallace 2017

Wallace BC, et al. Identifying reports of randomized controlled trials (RCTs) via a hybrid machine learning and crowdsourcing approach. J Am Med Inform Assoc 2017;24(6):1165-68.

Williamson 2011

Williamson I. Otitis media with effusion in children. BMJ Clinical Evidence 2011;2011:0502. [PMID: 19454116]

Zernotti 2017

Zernotti ME, Pawankar R, Ansotegui I, Badellino H, Croce JS, Hossny E, et al. Otitis media with effusion and atopy: is there a causal relationship? World Allergy Organization Journal 2017;10(1):37. [DOI: 10.1186/s40413-017-0168-x]

Zielhuis 1990

Zielhaus GA, Rach GH, van den Broek P. The occurrence of otitis media with effusion in Dutch pre-school children. Clinical Otolaryngology & Allied Sciences 1990;15(2):147-153.

Zimmermann 1992

Zimmerman IL, Steiner VG, Pond RE. Preschool Language Scale-3. San Antonio, TX: The Psychological Corporation, 1992.

References to other published versions of this review

MacKeith 2022b

MacKeith S, Mulvaney CA, Galbraith K, Marom T, Daniel M, Venekamp RP, et al. Ventilation tubes (grommets) for otitis media with effusion (OME) in children.

Figures and tables

Additional tables

Table 1

Sensitivity analyses

Outcome	Main analysis result (95% CI)	Sensitivity analysis	Sensitivity analysis result (95% CI)
Ventilation tubes versus no treatment			
Return to normal hearing			
1.1 Return to normal hearing, randomised by ear (medium-term)	OR 1.13 (0.46 to 2.74)	Correlation coefficient 0.3 instead of 0.5	OR 1.13 (0.46 to 2.74)
1.1 Return to normal hearing, randomised by ear (medium-term)	OR 1.13 (0.46 to 2.74)	Correlation coefficient 0.7 instead of 0.5	OR 1.13 (0.47 to 2.75)
1.1 Return to normal hearing, randomised by ear (medium-term)	OR 1.13 (0.46 to 2.74)	Normal hearing defined as <25dB HL instead of < 15 dB HL	OR 1.00 (0.57 to 1.76)
Final hearing threshold			
1.2 Mean final hearing threshold, randomised by ear (medium-term)	MD -3.47 (-9.97 to 3.03)	Correlation coefficient 0.3 instead of 0.5	MD -3.47 (-10.01 to 3.06)
1.2 Mean final hearing threshold, randomised by ear (medium-term)	MD -3.47 (-9.97 to 3.03)	Correlation coefficient 0.7 instead of 0.5	MD -3.49 (-10.37 to 3.38)
1.2 Mean final hearing threshold, randomised by ear (medium-term)	MD -3.47 (-9.97 to 3.03)	Fixed effect model	MD -3.31 (-5.09 to -1.54)
1.2 Mean final hearing threshold, randomised by ear (medium-term)	MD -3.47 (-9.97 to 3.03)	Exclusion of studies with concerns over trustworthiness	MD -9.90 (-13.00 to -6.80)
Change in hearing threshold from baseline			
1.3 Change in hearing threshold from baseline, randomised by ear (medium-term)	MD -0.16 (-3.28 to 2.97)	Correlation coefficient 0.3 instead of 0.5	MD -0.10 (-3.22 to 3.01)
1.3 Change in hearing threshold from baseline, randomised by ear (medium-term)	MD -0.16 (-3.28 to 2.97)	Correlation coefficient 0.7 instead of 0.5	MD -0.21 (-3.34 to 2.92)
Persistent tympanic membrane perforation			
1.4 Adverse event: perforation/retraction, randomised by ear (medium-term)	OR 0.85 (0.38 to 1.91)	Correlation coefficient 0.3 instead of 0.5	OR 0.85 (0.33 to 2.21)
1.4 Adverse event: perforation/retraction, randomised by ear (medium-term)	OR 0.85 (0.38 to 1.91)	Correlation coefficient 0.7 instead of 0.5	OR 0.91 (0.45 to 1.86)
1.4 Adverse event: perforation/retraction, randomised by ear (medium-term)	OR 0.85 (0.38 to 1.91)	Fixed effect model	OR 0.85 (0.38 to 1.91)
Persistence of OME			
1.6 Persistence of OME: randomised by child, analysed by ear (medium-term)	RR 0.30 (0.14 to 0.65)	Intracluster correlation of 1.0, instead of 0.5	RR 0.27 (0.11 to 0.70)
1.6 Persistence of OME: randomised by child, analysed by ear (medium-term)	RR 0.30 (0.14 to 0.65)	Intracluster correlation of 0, instead of 0.5	RR 0.30 (0.16 to 0.56)
1.7 Persistence of OME: randomised by ear (medium-term)	OR 0.66 (0.24 to 1.85)	Correlation coefficient 0.3 instead of 0.5	OR 0.66 (0.24 to 1.83)
1.7 Persistence of OME: randomised by ear (medium-term)	OR 0.66 (0.24 to 1.85)	Correlation coefficient 0.7 instead of 0.5	OR 0.66 (0.24 to 1.83)
1.7 Persistence of OME: randomised by ear (medium-term)	OR 0.66 (0.24 to 1.85)	Fixed effect model	OR 0.68 (0.42 to 1.09)
Ventilation tubes versus watchful waiting (treatment later if required)			
Final hearing threshold			
2.3 Mean final hearing threshold (air conduction), randomised by child (medium-term)	MD -1.89 (-7.32 to 3.54)	Fixed effect model	MD -0.74 (-3.08 to 1.59)
2.4 Mean final hearing threshold (air-bone gap), randomised by child, analysed by ear (medium-term)	MD -1.18 (-2.86 to 0.50)	Intracluster correlation of 1.0, instead of 0.5	MD -1.18 (-3.08 to 0.72)

2.4 Mean final hearing threshold (air-bone gap), randomised by child, analysed by ear (medium-term)	MD -1.18 (-2.86 to 0.50)	Intracluster correlation of 0, instead of 0.5	MD -1.18 (-2.58 to 0.22)
2.5 Mean final hearing threshold, randomised by child (long-term)	MD 0.36 (-0.41 to 1.13)	Correlation coefficient 0.3 instead of 0.5	MD 0.37 (-0.37 to 1.11)
2.5 Mean final hearing threshold, randomised by child (long-term)	MD 0.36 (-0.41 to 1.13)	Correlation coefficient 0.7 instead of 0.5	MD 0.35 (-0.45 to 1.16)
2.5 Mean final hearing threshold, randomised by child (long-term)	MD 0.36 (-0.41 to 1.13)	Fixed effect model	MD 0.36 (-0.41 to 1.13)
Persistent tympanic membrane perforation			
2.10 Adverse event: persistent perforation, randomised by child (long-term)	RR 3.65 (0.41 to 32.38)	Intracluster correlation of 1.0, instead of 0.5	RR 2.73 (0.29 to 25.97)
2.10 Adverse event: persistent perforation, randomised by child (long-term)	RR 3.65 (0.41 to 32.38)	Intracluster correlation of 0, instead of 0.5	RR 2.73 (0.56 to 13.43)
Persistence of OME			
2.11 Presence/persistence of OME, randomised by child, measured by otoscopy (medium-term)	RR 0.39 (0.09 to 1.72)	Intracluster correlation of 1.0, instead of 0.5	RR 0.49 (0.11 to 2.22)
2.11 Presence/persistence of OME, randomised by child, measured by otoscopy (medium-term)	RR 0.39 (0.09 to 1.72)	Intracluster correlation of 0, instead of 0.5	RR 0.40 (0.12 to 1.34)
2.14 Presence/persistence of OME, randomised by child (long-term)	RR 1.21 (0.84 to 1.74)	Fixed effect model	RR 1.22 (0.84 to 1.77)
Adverse events			
2.16 Adverse event: tympanosclerosis (long term)	RR 0.91 (0.33 to 2.55)	Intracluster correlation of 1.0, instead of 0.5	RR 0.91 (0.27 to 3.08)
2.16 Adverse event: tympanosclerosis (long term)	RR 0.91 (0.33 to 2.55)	Intracluster correlation of 0, instead of 0.5	RR 0.83 (0.36 to 1.92)
2.17 Adverse event: fibrosis (long term)	RR 0.61 (0.10 to 3.60)	Intracluster correlation of 1.0, instead of 0.5	RR 0.46 (0.04 to 4.97)
2.17 Adverse event: fibrosis (long term)	RR 0.61 (0.10 to 3.60)	Intracluster correlation of 0, instead of 0.5	RR 0.68 (0.15 to 3.03)
2.18 Adverse event: segmental atrophy (long term)	RR 2.83 (1.81 to 4.43)	Intracluster correlation of 1.0, instead of 0.5	RR 2.92 (1.72 to 4.96)
2.18 Adverse event: segmental atrophy (long term)	RR 2.83 (1.81 to 4.43)	Intracluster correlation of 0, instead of 0.5	RR 2.85 (1.97 to 4.13)
2.19 Adverse event: retraction pocket with other abnormality (long term)	RR 0.91 (0.06 to 14.41)	Intracluster correlation of 1.0, instead of 0.5	RR 0.91 (0.06 to 14.43)
2.19 Adverse event: retraction pocket with other abnormality (long term)	RR 0.91 (0.06 to 14.41)	Intracluster correlation of 0, instead of 0.5	RR 0.91 (0.06 to 14.64)
Psychosocial outcomes			
2.46 Parent-child interaction: Erickson child scale (medium-term)	MD -0.34 (-0.56 to -0.12)	Correlation coefficient 0.3 instead of 0.5 between five domains assessed	MD -0.34 (-0.53 to -0.15)
2.46 Parent-child interaction: Erickson child scale (medium-term)	MD -0.34 (-0.56 to -0.12)	Correlation coefficient 0.7 instead of 0.5 between five domains assessed	MD -0.34 (-0.58 to -0.10)
2.47 Parent-child interaction: Erickson parent scale (medium-term)	MD -0.42 (-0.67 to -0.17)	Correlation coefficient 0.3 instead of 0.5 between five domains assessed	MD -0.42 (-0.64 to -0.20)
2.47 Parent-child interaction: Erickson parent scale (medium-term)	MD -0.42 (-0.67 to -0.17)	Correlation coefficient 0.7 instead of 0.5 between five domains assessed	MD -0.42 (-0.70 to -0.14)
Ventilation tubes versus myringotomy			
Return to normal hearing			
4.1 Hearing returned to normal: VT versus laser myringotomy (medium-term)	RR 1.22 (0.59 to 2.53)	Intracluster correlation of 1.0, instead of 0.5	RR 1.21 (0.59 to 2.48)
4.1 Hearing returned to normal: VT versus laser myringotomy (medium-term)	RR 1.22 (0.59 to 2.53)	Intracluster correlation of 0, instead of 0.5	RR 1.22 (0.62 to 2.40)
4.1 Hearing returned to normal: VT versus laser myringotomy (medium-term)	RR 1.22 (0.59 to 2.53)	Fixed effect model	RR 1.33 (1.09 to 1.63)
4.1 Hearing returned to normal: VT versus laser myringotomy (medium-term)	RR 1.22 (0.59 to 2.53)	Exclusion of studies at high risk of bias	RR 1.00 (0.88 to 1.13)
Final hearing threshold			

4.2 Mean final hearing threshold, randomised by child (short-term).	RR 0.20 (-2.13 to 2.53)	Intracluster correlation of 1.0, instead of 0.5	RR 0.20 (-2.50 to 2.90)
4.2 Mean final hearing threshold, randomised by child (short-term).	RR 0.20 (-2.13 to 2.53)	Intracluster correlation of 0, instead of 0.5	RR 0.20 (-1.71 to 2.11)
4.4 Mean final hearing threshold (medium-term, pure tone audiometry)	MD 0.80 (-0.87 to 2.47)	Intracluster correlation of 1.0, instead of 0.5	MD 0.80 (-1.13 to 2.73)
4.4 Mean final hearing threshold (medium-term, pure tone audiometry)	MD 0.80 (-0.87 to 2.47)	Intracluster correlation of 0, instead of 0.5	MD 0.80 (-0.57 to 2.17)
Persistent tympanic membrane perforation			
4.5 Adverse event: persistent perforation (medium-term)	RR 1.00 (0.06 to 15.56)	Intracluster correlation of 1.0, instead of 0.5	RR 1.00 (0.06 to 15.45)
4.5 Adverse event: persistent perforation (medium-term)	RR 1.00 (0.06 to 15.56)	Intracluster correlation of 0, instead of 0.5	RR 2.00 (0.19 to 21.54)
4.6 Adverse event: persistent perforation cold-steel myringotomy (medium-term)	Peto OR 8.09 (1.78 to 36.79)	Exclusion of studies with concerns over trustworthiness	Peto OR 7.39 (0.15 to 372.38)
Persistence of OME			
4.7 Persistence of OME: VT versus laser myringotomy (short-term)	RR 1.40 (0.48 to 4.12)	Intracluster correlation of 1.0, instead of 0.5	RR 1.50 (0.46 to 4.92)
4.7 Persistence of OME: VT versus laser myringotomy (short-term)	RR 1.40 (0.48 to 4.12)	Intracluster correlation of 0, instead of 0.5	RR 1.43 (0.58 to 3.53)
4.10 Persistence of OME: VT versus laser myringotomy (medium-term)	RR 0.32 (0.16 to 0.64)	Intracluster correlation of 1.0, instead of 0.5	RR 0.35 (0.17 to 0.74)
4.10 Persistence of OME: VT versus laser myringotomy (medium-term)	RR 0.32 (0.16 to 0.64)	Intracluster correlation of 0, instead of 0.5	RR 0.33 (0.18 to 0.60)
4.11 Persistence of OME: VT versus laser myringotomy, randomised by ear (medium-term)	OR 0.27 (0.19 to 0.38)	Correlation coefficient 0.3 instead of 0.5	OR 0.27 (0.18 to 0.42)
4.11 Persistence of OME: VT versus laser myringotomy, randomised by ear (medium-term)	OR 0.27 (0.19 to 0.38)	Correlation coefficient 0.7 instead of 0.5	OR 0.27 (0.21 to 0.36)
Adverse events			
4.20 Adverse event: retraction of TM: VT versus laser myringotomy (medium-term)	RR 2.67 (0.75 to 9.48)	Intracluster correlation of 1.0, instead of 0.5	RR 3.50 (0.77 to 15.85)
4.20 Adverse event: retraction of TM: VT versus laser myringotomy (medium-term)	RR 2.67 (0.75 to 9.48)	Intracluster correlation of 0, instead of 0.5	RR 2.75 (0.92 to 8.21)
4.22 Adverse event: otorrhoea: VT versus laser myringotomy (medium-term)	RR 4.00 (0.46 to 34.57)	Intracluster correlation of 1.0, instead of 0.5	RR 3.00 (0.33 to 27.66)
4.22 Adverse event: otorrhoea: VT versus laser myringotomy (medium-term)	RR 4.00 (0.46 to 34.57)	Intracluster correlation of 0, instead of 0.5	RR 2.50 (0.50 to 12.44)

CI confidence interval; MD mean difference; OR odds ratio; RR risk ratio

Table 2

Study features

Study	Participants	Setting	Intervention	Comparator	Concomitant treatment	Follow-up (main outcomes reported at this time)	No
Bernard 1991	Children aged 2.5 to 7 years with OME and unsuccessful treatment with 2 courses of antibiotics (n = 139)	Single centre, USA	Bilateral myringotomy and insertion of ventilation tubes	Antibiotics (Sulfisoxazole, 75mg/kg divided into 2 daily doses for 6 months)	None reported	18 months	
D'Eredita 2006	Children aged 2-6 with OME (n = 30)	Single centre, Italy	Cold myringotomy and ventilation tube insertion (unclear if bilateral or unilateral)	Laser myringotomy	Ofloxacin solution three times daily for 5 days	12 months	

Dempster 1993	Children aged 3.5 to 12 years with bilateral OME (n = 78)	Single centre, UK	Unilateral ventilation tube	No ventilation tube	Half of the children in this study also underwent adenoidectomy.	11 months	Childre receive ventilat in one e no treat the othe
Elkholy 2021	Children aged 5-15 years with OME (n = 40)	Single centre, Egypt	Ventilation tube insertion (unclear if bilateral or unilateral)	No treatment	Children also underwent adenoidectomy	2 weeks	Additio follow-u months useable were re after 2
Gates 1989	Children aged 4-8 years with persistent OME for 60 days after a 10-day course of erythromycin and sulfisoxazole, and a 30-day course of pseudoephedrine hydrochloride (n = 578)	Multicentre, USA	Bilateral ventilation tubes or Adenoidectomy plus bilateral ventilation tubes	Myringotomy or Adenoidectomy plus myringotomy		2 years	4-arm t
Koopman 2004	Children aged <11 years with bilateral OME (n = 208)	Multicentre, Netherlands	Ventilation tube	Laser myringotomy		6 months	Childre receive interve each ea
Maw 1983	Children aged 2-9 years with bilateral OME (n = 145)	Single centre, UK	Ventilation tubes	No treatment	Half of the children in this study also underwent adenoidectomy.	3 years	
Maw 1999	Children aged 9 months to 4.5 years with bilateral OME (n = 182)	Single centre, UK	Bilateral ventilation tubes	Watchful waiting		Up to 7 years	21% of particip the wat waiting receive surgery 9 montl months 85% of particip this gro been lis or alrea receive surgery
Paradise 2007	Children aged <3 years with OME (n = 429)	Multicentre, USA	Ventilation tubes	Watchful waiting		Up to 11 years	45% of the wat waiting had rec ventilat by the a 11 year
Popova 2010	Children (mean age 5 years) with bilateral OME (n = 90)	Single centre, Bulgaria	Ventilation tubes	Myringotomy	All participants received adenoidectomy	12 months	
Rach 1991	Children aged 2-4 years with OME (n = 43)	Single centre, Netherlands	Ventilation tubes	No treatment		4 years	After 6 some c in the 'r treatme underw insertio therefo from lat points a include

							compar VT with waiting
Rovers 2000	Children (mean age 19.5 months) who failed three successive hearing tests with bilateral OME (n = 187)	Multicentre, Netherlands	Ventilation tubes	Watchful waiting		12 months	
Ruckley 1988	Children aged 4-9 years with bilateral OME (n = 40)	Single centre, UK	Ventilation tube	Thermal myringotomy	Adenoidectomy	3 months	Childre receive interver each ea
Sujatha 2015	Children aged 3-10 years with OME (n = 50)	Single centre, India	Ventilation tube	Myringotomy	Adenoidectomy. Systemic antibiotics, analgesics, anti-inflammatory and decongestant nasal drops for 7 days.	12 months	
Tao 2020	Children aged 4-12 years with bilateral OME (n = 178)	Single centre, China	Ventilation tube	Myringotomy	Adenoidectomy	12 months	
TARGET 2000	Children aged 3.25 to 6.75 with bilateral OME (n = 248)	Multicentre, UK	Bilateral ventilation tubes alone	Watchful waiting		2 years	Additio arm inc the con review adenoi
To 1984	Children aged <14 years with bilateral OME (n = 54)	Single centre, UK	Ventilation tube	Myringotomy	Adenoidectomy	1-5 years	Childre receive interver each ea
Velepic 2011	Children (mean age 5.5. years) with predominantly bilateral OME (n = 87)	Single centre, Croatia	Ventilation tube	Watchful waiting (ventilation tube after 3 months if required)	Adenoidectomy	6 months	
Yousaf 2016	Children aged 4-12 years with OME and hearing level >30db HL (n = 82)	Single centre, Pakistan	Ventilation tube	Laser myringotomy		6 months	

Table 3

Adverse events: primary and secondary outcomes: tympanic membrane changes and tube related

Comparison and studies	Primary outcome	Secondary outcomes				
		1. Tympanic membrane changes				
	Persistent perforation	Tympanosclerosis	Myringosclerosis	Infection	Foreign body reaction	Other
VT vs no treatment						
Dempster 1993	VT: 6/72 (8.3%) No VT: 7/72 (9.7%) (described as persistent	VT: 28/72 (39%) No VT: 1/72 (1.4%)	x	x	x	x

	perforation or retraction)					
Maw 1983	x	x	x	x	x	x
Rach 1991	x	x	x	x	x	x
Early VT vs Watchful waiting (treatment later if required)						
TARGET 2000	VT (with and without adenoidectomy) : 8/635 (0.01%) =/> 6 months see Effects of interventions	VT (with and without adenoidectomy) : 128/635 (20%) WW 0 =/> 6 months see Effects of interventions	x	x	x	x
Maw 1999	x	x	x	x	x	x
Paradise 2007		see Analysis 2.16				see Analysis 2.17; Analys 2.18; Analys 2.19
Rach 1991 (long term data)	x	x	x	x	x	x
Rovers 2000	x	x	x	x	x	x

Velepic 2011	see Analysis 2.9	x	Total of 42/161 (26%) ears see Effects of interventions	x	x	1. Attic retraction: total of 79/161 (49% ears showe attic retraction 2. Tensa retractions with/without malleus rotation tota of 36/161 (22%) ears 3. Scars of the ear drun total of 46/161 (29%) see Effects interventions
VT vs non-surgical treatment						
Bernard 1991	VT: 0/60 (0%) 18 months	x	VT: 17/60 (28.3%) Antibiotic 4/65 (6.1%) 18 months	VT: 17/60 (28.3%) 18 months	VT: 17/60 (28.3%) 18 months	x
VT vs myringotomy alone						
ACTRN12611001073998	no data available as yet					
D'Eredita 2006	VT: 1/15 (6.7%) LM: no data reported for LM	x	X	X	X	x
Gates 1989	In 6 children (3 post myringotomy and 3 post VT (group allocations not reported). see Effects of interventions		X	X	X	

Koopman 2004	X	x	X	X	X	x
Popova 2010	x	x	x	x	x	x
Ruckley 1988	x	VT 0/36 (0%), TM 0/36 (0%0 short term 3 mo	x	x	x	x
Sujatha 2015	VT: R ear: 3/25 L ear: 3/25 Myringotomy R ear: 0/25	R ear 12 months: Tympanosclerotic patch Myringotomy 1/25 (4%)	x	x	x	R ear: 3 months: Retraction Myringotom 22/25 (88%)

	L ear: 0/25 12 months	VT 2/25 (8%) L ear: 12 months: Tympanosclerotic patch Myringotomy 0/25 VT 1/25 (4%)				12 months: Retraction Myringotom 7/25 (28%) VT 14/25 (56%) L ear: 3 months: Retraction Myringotom 22/25 (88%) 12 months: Retraction Myringotom 6/25 (24%) VT 12/25 (48%)
Tao 2020	VT: 12 months: 4 ears/4 patients	VT: 12 months: 6 ears/5 patients (calcified plaques)	x	x	x	x
To 1984	Authors state "One ear which had received a grommet was improving but was still abnormal." Presumed 1/54 for VT.	VT 9/54, Myringotomy 1/54 timing of follow-up not reported	x	x	x	Retraction segments V 2/54, Myringotom 1/54 24 month long term
Yousaf 2016		x	x	x	x	

	see Analysis 4.5					1. Hypertrophi scar, see Analysis 4.18 2. Retractio of tympanic membrane, see Analysis 4.17

Table 4

Adverse events: secondary outcomes: patient related

Comparison and studies	Secondary outcomes					
	3. Patient related					
	Serious medication-related side effects	Allergic reaction (appearing within 7 days of starting treatment)	Nausea	Vomiting	Otalgia	Post surgical haemorrhage
VT vs no treatment						
Dempster 1993	x	x	x	x	x	x
Maw 1983						
Rach 1991	x	x	x	x	x	x
Early VT vs Watchful Waiting (treatment later if required)						
TARGET 2000	x	x	x	x	x	1/165 (0.6%) children that had adenoidectomy had to return to theatre for postoperative haemorrhage. (*Note: N exceeds number allocated to Ad group because of cross-overs from other groups*)
Maw 1999	x	x	x	x	x	x
Paradise 2007	x	x	x	x	x	x
Rach 1991 (long term data only)	x	x	x	x	x	x
Rovers 2000	x	x	x	x	x	x
Velepic 2011	x	x	x	x	x	x
VT vs non-surgical treatment						
Bernard 1991	Sulfonamide: 0/65 (0%) 18 mo	Sulfonamide: 4/65 (6.2%) 18 mo	Sulfonamide: 2/65 (3.1%) 18 mo	Sulfonamide: 0/65 (0%) 18 mo	X	X
VT vs myringotomy alone						
ACTRN12611001073998	no data available as yet					
D'Eredita 2006	X	X	X	X	X	X
Gates 1989	X	X	X	X	X	1/251 after adenoidectomy (unclear why 251). Returned to Operating theatre for control
Koopman 2004	X	X	X	X	LM 1/208 (0.5%)	X

					during first 2 days post LM	
Popova 2010	x	x	x	x	x	x
Ruckley 1988	x	x	x	x	TM 1/36 (2.8%) VT not reported v short term	x
Sujatha 2015	x	x	x	x	x	x
Tao 2020	x	x	x	x	x	x
To 1984	x	x	x	x	x	x
Yousaf 2016	x	x	x	x	x	LM 0, VT 9 (13%)

Table 5

Developmental outcomes at age 9 to 11 from Paradise 2007 with GRADE of certainty

Test	Reported test properties, working MID	Early VT mean score ±SD (n)	WW mean score ±SD (n)	MD (95%CI)	GRADE of certainty[a]
Literacy					
Woodcock Reading Mastery Tests:	The normative mean standard score is 100±15. Higher scores indicate more favourable results. Working MID of 15.				
Word identification subtest		98±11 (195)	99±12 (196)	-1.00 (-3.28, 1.28)	Very low
Word Attack subtest		103±13 (195)	104±14 (196)	-1.00 (-3.68 to 1.68)	Very low
Passage Comprehension subtest		98±12 (195)	99±12 (196)	-1.00 (-3.38, 1.38)	Very low
Oral reading fluency test:	Higher scores indicate more favourable results. Working MID of 15.				
Children in grade 3		78±36 (37)	87±41 (37)	-9.00 (-26.58 to 8.58)	Very low
Children in grade 4		89±36 (87)	89±38 (97)	0.00 (-10.70 to 10.701)	Very low
Children in grade 5		97±36 (54)	102±37 (51)	-5.00 (-18.98 to 8.98)	Very low
Children in grade 6		102±32 (12)	96±43 (9)	6.00 (-27.42, 39.42)	Very low
Woodcock–Johnson III Tests of Achievement:	In both subtests, raw scores are converted to standard scores according to the child's age. The normative mean standard score on both subtests is 100±15. Higher scores indicate more favourable results. Working MID of 15.				
Spelling subtest		96±13 (194)	97±16 (196)	-1.00 (-3.89 to 1.89)	Very low
Writing Samples subtest		104±14 (192)	105±15 (195)	-1.00 (-3.89 to 1.89)	Very low
Phonological awareness					

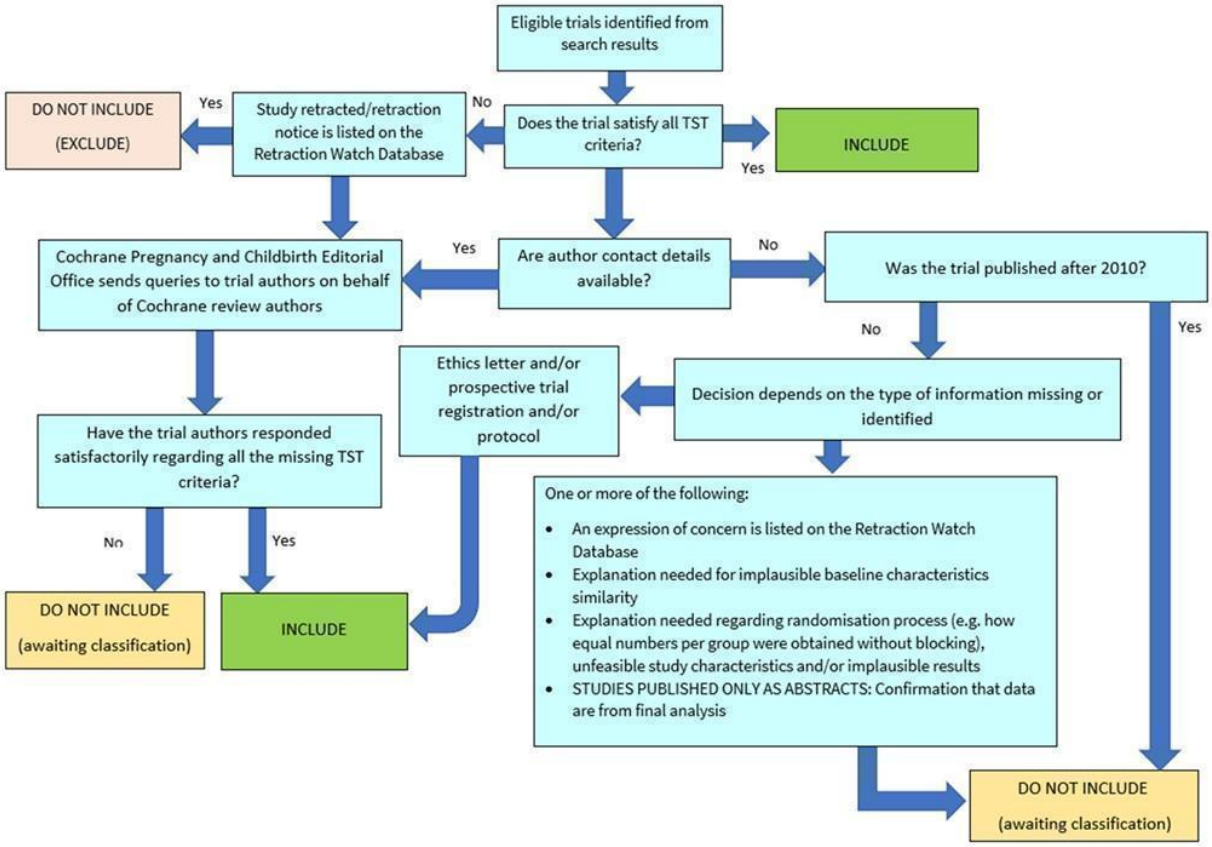
Comprehensive Test of Phonological Processing:	In both subtests, raw scores are converted to standard scores according to the child's age. The normative mean standard score on each subtest is 10±3. Higher scores indicate more favourable results. Working MID of 3.				
Elision subtest		8.6±4.9 (195)	8.7±3.0 (196)	-0.10 (-0.91 to 0.71)	Very low
Rapid Letter Naming subtest		9.3±2.5 (193)	9.6±2.4 (196)	-0.30 (-0.79 to 0.19)	Very low
Attention, impulsivity, and psychosocial function					
Disruptive Behavior Disorders Rating Scale	The items are scored on a four-point scale (0, "not at all"; 1, "just a little"; 2, "pretty much; 3, "very much) and are averaged for comparison with normative data. For boys 9 or 10 years of age, the normative mean score for the inattention factor is 1.01±0.91; for the impulsivity and overactivity factor, 0.86±0.81; and for the oppositional defiant factor, 0.69±0.77. For boys 11 through 14 years of age, the corresponding values are 1.01±0.96, 0.85±0.88, and 0.73±0.86. Normative data for girls are not available. Higher scores indicate less favourable results. Working MID of 0.96 (inattention), 0.88 (impulsivity and overactivity) and 0.86 (oppositional defiant factor).				
Inattention factor:					
Parent's rating		0.70±0.63 (194)	0.65±0.66 (196)	0.05 (-0.08 to 0.18)	Very low
Teacher's rating		0.71±0.74 (190)	0.67±0.75 (192)	0.04 (-0.11 to 0.19)	Very low
Impulsivity and overactivity factor:					
Parent's rating		0.67±0.57 (194)	0.57±0.54 (196)	0.10 (-0.01 to 0.21)	Very low
Teacher's rating		0.48±0.63 (190)	0.40±0.52 (192)	0.08 (-0.04 to 0.20)	Very low
Oppositional defiant factor:					
Parent's rating		0.57±0.58 (194)	0.52±0.53 (196)	0.05 (-0.06 to 0.16)	Very low
Teacher's rating		0.33±0.56 (190)	0.33±0.58 (192)	0.00 (-0.11 to 0.11)	Very low
Child Behavior Checklist:	Scores on each of the eight component scales and a Total Problem score are calculated and converted to T scores. The normative mean T score on each scale and for Total Problems is 50±10. Only the Total Problem scores are shown here. Higher scores indicate less favourable results. Working MID of 10.				
Total Problems score, parent's rating		51±12 (194)	49±12 (196)	2.00 (-0.38, 4.38)	Very low
Total Problems score, teacher's rating		52±11 (189)	50±11 (191)	2.00 (-0.21 to 4.21)	Very low

Impairment Rating Scales:	A score of 3 or higher is considered to be indicative of clinically meaningful impairment. Working MID of 3.				
Overall functioning, parent's rating		0.82±1.42 (194)	0.68±1.33 (196)	0.14 (-0.13 to 0.41)	Very low
Overall functioning, teacher's rating		2.04±2.24 (190)	1.78±2.19 (192)	0.26 (-0.18, 0.70)	Very low
Social Skills Rating System:	The normative mean standard score is 100±15. Higher scores indicate more favourable results. Working MID of 15.				
Social Skills scale, parent's version		96±19 (194)	98±18 (194)	-2.00 (-5.68 to 1.68)	Very low
Social Skills scale, teacher's version		98±13 (184)	99±13 (186)	-1.00 (-3.65 to 1.65)	Very low
Visual Continuous Performance Test:	Normative data are not available. Higher scores indicate less favourable results. Working MID of 2.				
Inattention		9.7±8.5 (195)	9.5±8.5 (196)	0.20 (-1.49 to 1.89)	Very low
Impulsivity		8.8±16.5 (195)	8.2±15.6 (196)	0.60 (-2.58 to 3.78)	Very low
Auditory Continuous Performance Test:	Normative data are not available. Higher scores indicate less favourable results. Working MID of 2.				
Inattention		11.1±7.2 (155)	11.4±8.0 (153)	-0.30 (-2.00 to 1.40)	Very low
Impulsivity		3.3±8.7 (154)	4.2±12.1 (153)	-0.90 (-3.26 to 1.46)	Very low
Intelligence and academic achievement					
Wechsler Abbreviated Scale of Intelligence	The normative mean score is 100±15. Higher scores indicate more favourable results. Working MID of 15.	96±13 (195)	96±14 (196)	0.00 (-2.68 to 2.68)	Very low
Calculation subtest of the Woodcock–Johnson III Tests of Achievement	The normative mean score is 100±15. Higher scores indicate more favourable results. Working MID of 15.	99±13 (194)	99±13 (195)	0.00 (-2.58, 2.58)	Very low

VT: ventilation tubes; WW: watchful waiting; MD: mean difference; MID: minimum important difference.

a: GRADING for risk of bias, inconsistency, indirectness and publication bias was the same for each effect estimate (downgraded two levels for performance bias, no downgrade, downgraded one level for population indirectness and no downgrade respectively). Imprecision was downgraded one level for each effect estimate as the optimal information size was not attained, and downgraded a further level when two decision thresholds were crossed by the CI.

Figure 1



The Cochrane Pregnancy and Childbirth Trustworthiness Screening Tool

Figure 2

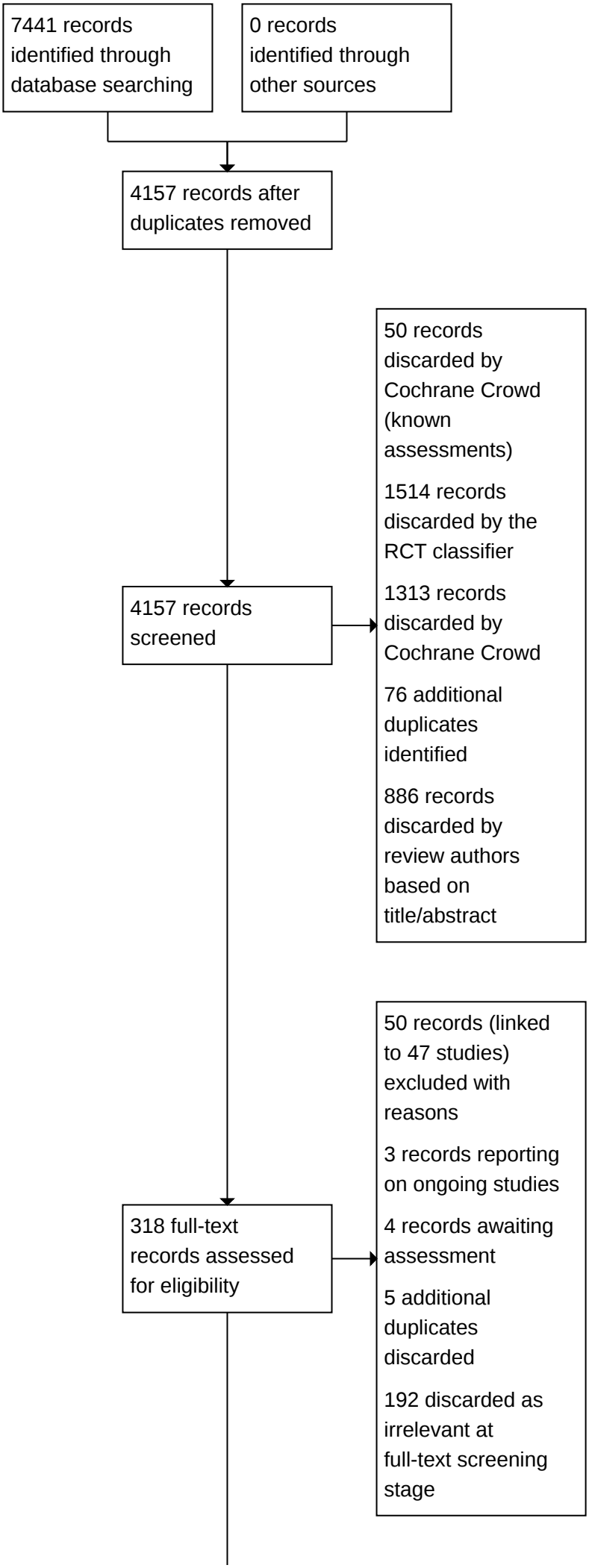

7441 records identified through database searching
0 records identified through other sources
4157 records after duplicates removed
4157 records screened
50 records discarded by Cochrane Crowd (known assessments)
1514 records discarded by the RCT classifier
1313 records discarded by Cochrane Crowd
76 additional duplicates identified
886 records discarded by review authors based on title/abstract
318 full-text records assessed for eligibility
50 records (linked to 47 studies) excluded with reasons
3 records reporting on ongoing studies
4 records awaiting assessment
5 additional duplicates discarded
192 discarded as irrelevant at full-text screening stage

Figure 3

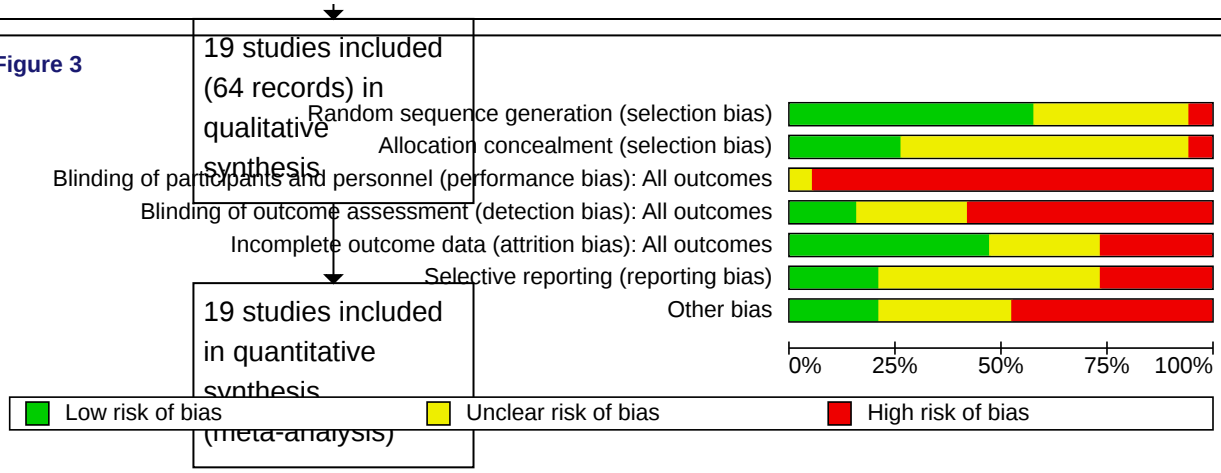
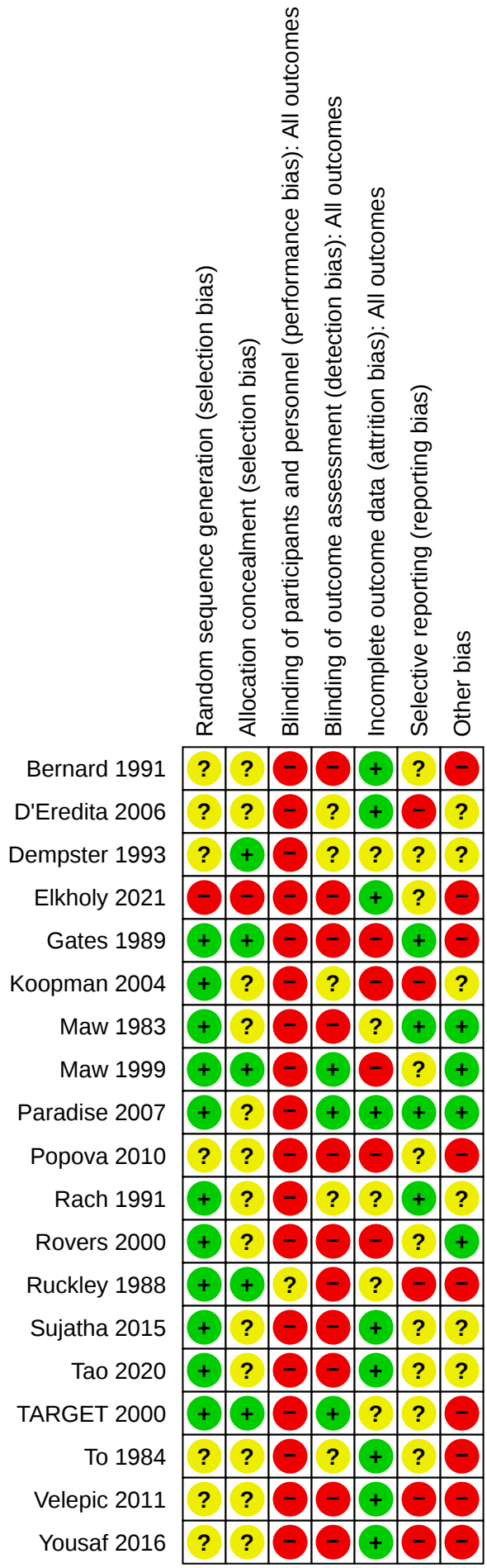
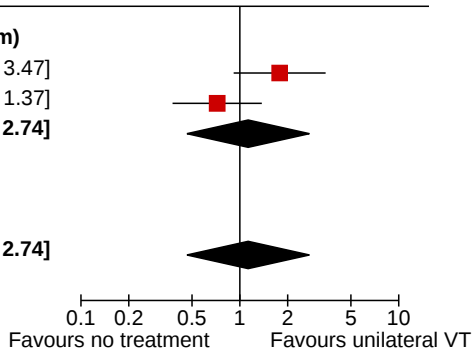


Figure 4


Random sequence generation (selection bias)
Allocation concealment (selection bias)
Blinding of participants and personnel (performance bias): All outcomes
Blinding of outcome assessment (detection bias): All outcomes
Incomplete outcome data (attrition bias): All outcomes
Selective reporting (reporting bias)
Other bias
Bernard 1991
D'Eredita 2006
Dempster 1993
Elkholy 2021
Gates 1989
Koopman 2004
Maw 1983
Maw 1999
Paradise 2007
Popova 2010
Rach 1991
Rovers 2000
Ruckley 1988
Sujatha 2015
Tao 2020
TARGET 2000
To 1984
Velepic 2011
Yousaf 2016

Analysis 1.1

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
1.1.1 Randomised by ear: normal defined as <15dB. CC=0.5 (medium term)				
Dempster 1993 (1)	0.576613	0.34	49.6%	1.78 [0.91 , 3.47]
Dempster 1993 (2)	-0.328504	0.33	50.4%	0.72 [0.38 , 1.37]
Subtotal (95% CI)			**100.0%**	**1.13 [0.46 , 2.74]**
Heterogeneity: Tau² = 0.30; Chi² = 3.65, df = 1 (P = 0.06); I² = 73%				
Test for overall effect: Z = 0.27 (P = 0.79)				
Total (95% CI)			**100.0%**	**1.13 [0.46 , 2.74]**
Heterogeneity: Tau² = 0.30; Chi² = 3.65, df = 1 (P = 0.06); I² = 73%				
Test for overall effect: Z = 0.27 (P = 0.79)				
Test for subgroup differences: Not applicable				



Footnotes

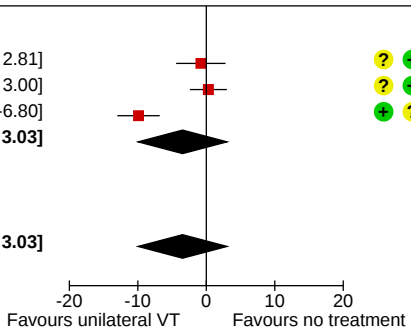
(1) Unilateral VT versus no treatment at 12 months

(2) Adenoidectomy and unilateral VT versus adenoidectomy only at 12 months

Comparison 1: VT versus no treatment, Outcome 1: Return to normal hearing, randomised by ear (medium-term)

Analysis 1.2

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
1.2.1 Correlation coefficient = 0.5					
Dempster 1993 (1)	-0.8	1.84	32.5%	-0.80 [-4.41 , 2.81]	? + − ? ? ? ?
Dempster 1993 (2)	0.3	1.38	34.0%	0.30 [-2.40 , 3.00]	? + − ? ? ? ?
Maw 1983 (1)	-9.9	1.58	33.4%	-9.90 [-13.00 , -6.80]	+ ? − − ? + +
Subtotal (95% CI)			**100.0%**	**-3.47 [-9.97 , 3.03]**	
Heterogeneity: Tau² = 30.42; Chi² = 26.10, df = 2 (P < 0.00001); I² = 92%					
Test for overall effect: Z = 1.05 (P = 0.30)					
Total (95% CI)			**100.0%**	**-3.47 [-9.97 , 3.03]**	
Heterogeneity: Tau² = 30.42; Chi² = 26.10, df = 2 (P < 0.00001); I² = 92%					
Test for overall effect: Z = 1.05 (P = 0.30)					
Test for subgroup differences: Not applicable					



Footnotes

(1) Unilateral VT versus nil at 12 months

(2) Ad + unilateral VT versus ad only at 12 months

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

(E) Incomplete outcome data (attrition bias)

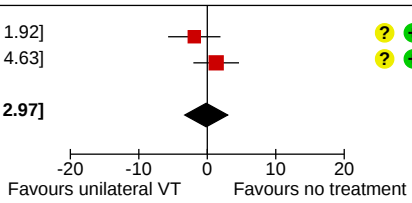
(F) Selective reporting (reporting bias)

(G) Other bias

Comparison 1: VT versus no treatment, Outcome 2: Mean final hearing threshold, randomised by ear (medium-term)

Analysis 1.3

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Dempster 1993 (1)	-1.9	1.95	45.5%	-1.90 [-5.72 , 1.92]		? + − ? ? ? ?
Dempster 1993 (2)	1.3	1.7	54.5%	1.30 [-2.03 , 4.63]		? + − ? ? ? ?
Total (95% CI)			**100.0%**	**-0.16 [-3.28 , 2.97]**		



Heterogeneity: Tau² = 1.77; Chi² = 1.53, df = 1 (P = 0.22); I² = 35%
Test for overall effect: Z = 0.10 (P = 0.92)
Test for subgroup differences: Not applicable

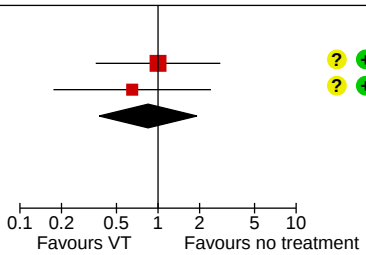
Footnotes
(1) Unilateral VT versus nil. CC=0.5
(2) Adenoidectomy plus unilateral VT versus adenoidectomy only. CC=0.5

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 1: VT versus no treatment, Outcome 3: Change in hearing threshold from baseline, randomised by ear (medium-term)

Analysis 1.4

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI	Odds Ratio IV, Random, 95% CI	Risk of Bias A B C D E F G
1.4.1 Correlation coefficient 0.5						
Dempster 1993 (1)	0	0.53	61.5%	1.00 [0.35 , 2.83]		? + − ? ? ? ?
Dempster 1993 (2)	-0.430783	0.67	38.5%	0.65 [0.17 , 2.42]		? + − ? ? ? ?
Subtotal (95% CI)			**100.0%**	**0.85 [0.38 , 1.91]**		



Heterogeneity: Tau² = 0.00; Chi² = 0.25, df = 1 (P = 0.61); I² = 0%
Test for overall effect: Z = 0.40 (P = 0.69)

Test for subgroup differences: Not applicable

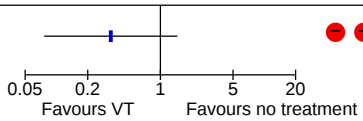
Footnotes
(1) Unilateral VT + Ad versus Ad only, at 12 months.
(2) Unilateral VT versus no treatment at 12 months.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 1: VT versus no treatment, Outcome 4: Adverse event: perforation/retraction, randomised by ear (medium-term)

Analysis 1.5

Study or Subgroup	VT Events	VT Total	No treatment Events	No treatment Total	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
Elkholy 2021 (1)	2	20	6	20	0.33 [0.08 , 1.46]		− − − − + ? −



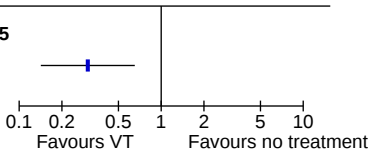
Footnotes
(1) 2 weeks follow-up.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 1: VT versus no treatment, Outcome 5: Persistence of OME: randomised by child (very short-term)

Analysis 1.6

Study or Subgroup	VT Events	VT Total	No treatment Events	No treatment Total	Risk Ratio M-H, Random, 95% CI
1.6.1 Adjusted for non-independence of within-individual measurements, assuming ICC of 0.5					
Rach 1991 (1)	6	29	17	25	0.30 [0.14 , 0.65]



Footnotes

(1) Bilateral VT versus nil at 6 months. Analysed by ear. Average cluster size=2. DE=1.5

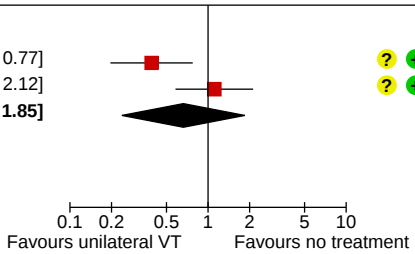
Comparison 1: VT versus no treatment, Outcome 6: Persistence of OME: randomised by child (medium-term)

Analysis 1.7

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
1.7.1 Correlation coefficient = 0.5				
Dempster 1993 (1)	-0.941609	0.35	49.4%	0.39 [0.20 , 0.77]
Dempster 1993 (2)	0.10436	0.33	50.6%	1.11 [0.58 , 2.12]
Subtotal (95% CI)			**100.0%**	**0.66 [0.24 , 1.85]**

Heterogeneity: Tau² = 0.43; Chi² = 4.73, df = 1 (P = 0.03); I² = 79%

Test for overall effect: Z = 0.79 (P = 0.43)



Footnotes

(1) Unilateral VT versus nil. Tympanometry at 12 months.

(2) Adenoidectomy plus unilateral VT versus adenoidectomy alone

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

(E) Incomplete outcome data (attrition bias)

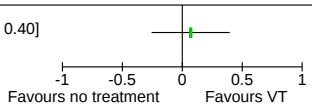
(F) Selective reporting (reporting bias)

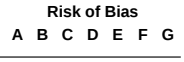
(G) Other bias

Comparison 1: VT versus no treatment, Outcome 7: Persistence of OME: randomised by ear (medium-term)

Analysis 1.8

Study or Subgroup	VT Mean	VT SD	VT Total	No treatment Mean	No treatment SD	No treatment Total	Mean Difference IV, Random, 95% CI
Rach 1991 (1)	0.17	0.563857	22	0.1	0.527247	21	0.07 [-0.26 , 0.40]





Footnotes

(1) Comparison of mean improvement in z-score on Reynell test for verbal comprehension.

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

(E) Incomplete outcome data (attrition bias)

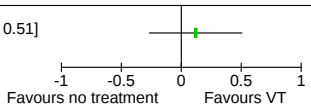
(F) Selective reporting (reporting bias)

(G) Other bias

Comparison 1: VT versus no treatment, Outcome 8: Mean improvement in comprehensive language, randomised by child (medium-term)

Analysis 1.9

Study or Subgroup	VT Mean	VT SD	VT Total	No treatment Mean	No treatment SD	No treatment Total	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Rach 1991 (1)	0.29	0.681027	21	0.17	0.587589	20	0.12 [-0.27 , 0.51]		+ ? − ? ? + ?



Footnotes

(1) Comparison of mean improvement in z-score on Reynell test for verbal comprehension.

Risk of bias legend

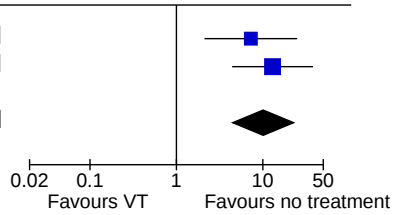
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 1: VT versus no treatment, Outcome 9: Mean improvement in expressive language, randomised by child (medium-term)

Analysis 1.10

Study or Subgroup	VT Events	VT Total	No treatment Events	No treatment Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI	Peto Odds Ratio Peto, Fixed, 95% CI
Dempster 1993 (1)	11	35	1	35	43.2%	7.26 [2.11 , 24.95]	
Dempster 1993 (2)	17	37	0	37	56.8%	12.95 [4.42 , 37.99]	
Total (95% CI)		**72**		**72**	**100.0%**	**10.09 [4.48 , 22.70]**	
Total events:	28		1				

Heterogeneity: Chi² = 0.48, df = 1 (P = 0.49); I² = 0%
Test for overall effect: Z = 5.59 (P < 0.00001)
Test for subgroup differences: Not applicable



Footnotes

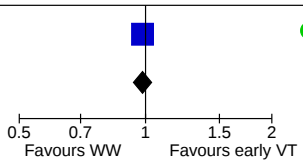
(1) Unilateral VT versus no treatment at 12 months.
(2) Unilateral VT + Ad versus Ad only, at 12 months.

Comparison 1: VT versus no treatment, Outcome 10: Adverse event: tympanosclerosis, randomised by ear (medium-term)

Analysis 2.1

Study or Subgroup	Early VT Events	Early VT Total	Watchful waiting Events	Watchful waiting Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
Paradise 2007 (1)	182	195	186	196	100.0%	0.98 [0.94 , 1.03]		+ ? − + + + +
Total (95% CI)		**195**		**196**	**100.0%**	**0.98 [0.94 , 1.03]**		
Total events:	182		186					

Heterogeneity: Not applicable
Test for overall effect: Z = 0.66 (P = 0.51)
Test for subgroup differences: Not applicable



Footnotes

(1) Age 9 to 11. Hearing-level threshold of 15 dB HL or less at 1000, 2000, and 4000 Hz.

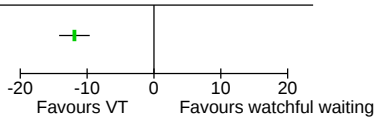
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 1: Hearing returned to normal, randomised by child (long-term)

Analysis 2.2

Study or Subgroup	VT Mean	VT SD	VT Total	Watchful waiting Mean	Watchful waiting SD	Watchful waiting Total	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI
TARGET 2000 (1)	14.4	6.9	109	26.3	9.9	106	-11.90 [-14.19 , -9.61]	



Footnotes

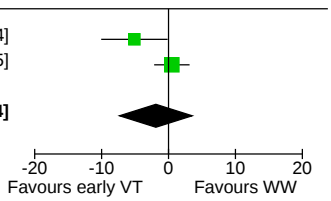
(1) Bilateral VT versus WW at 3 months. Maximum cases available.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 2: Mean final hearing threshold, randomised by child (short-term)

Analysis 2.3

Study or Subgroup	Early VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Maw 1999 (1)	16.5	13	81	21.6	16.1	60	42.7%	-5.10 [-10.06 , -0.14]
TARGET 2000 (2)	21	9.4	110	20.5	10.1	100	57.3%	0.50 [-2.15 , 3.15]
Total (95% CI)			**191**			**160**	**100.0%**	**-1.89 [-7.32 , 3.54]**

Heterogeneity: Tau² = 11.57; Chi² = 3.81, df = 1 (P = 0.05); I² = 74%
Test for overall effect: Z = 0.68 (P = 0.49)
Test for subgroup differences: Not applicable

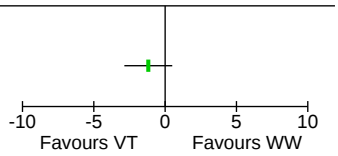


Footnotes
(1) Bilateral VT versus WW at 9 months, best ear at 4000Hz.
(2) Bilateral VT versus WW at 12 months. Maximum cases available.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 3: Mean final hearing threshold (air conduction), randomised by child (medium-term)

Analysis 2.4

Study or Subgroup	VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Mean Difference IV, Random, 95% CI
2.4.1 Adjusted for non-independence of within-individual measurements, assuming ICC of 0.5							
Velepic 2011 (1)	6.02	3.81	41	7.2	5.19	71	-1.18 [-2.86 , 0.50]



Footnotes
(1) VT + Ad versus WW + Ad. Reported by ear, PTA air-bone gap. Ave cluster size-=1.85. DE=1.425

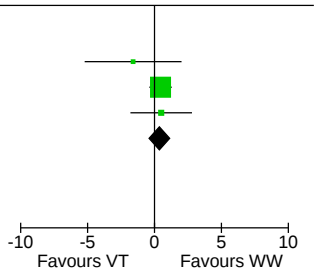
Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 4: Mean final hearing threshold (air-bone gap), randomised by child, analysed by ear (medium-term)

Analysis 2.5

Study or Subgroup	VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.5.1 Assume correlation coefficient for Paradise 2007 (left and right ear data combined) of 0.5								
Maw 1999 (1)	12.7	11.5	75	14.3	10.5	67	4.5%	-1.60 [-5.22 , 2.02]
Paradise 2007 (2)	6.2	3.55	147	5.75	3.6	134	84.3%	0.45 [-0.39 , 1.29]
TARGET 2000 (3)	18.7	8.9	108	18.2	8.1	102	11.2%	0.50 [-1.80 , 2.80]
Subtotal (95% CI)			**330**			**303**	**100.0%**	**0.36 [-0.41 , 1.13]**

Heterogeneity: Tau² = 0.00; Chi² = 1.19, df = 2 (P = 0.55); I² = 0%
Test for overall effect: Z = 0.93 (P = 0.35)

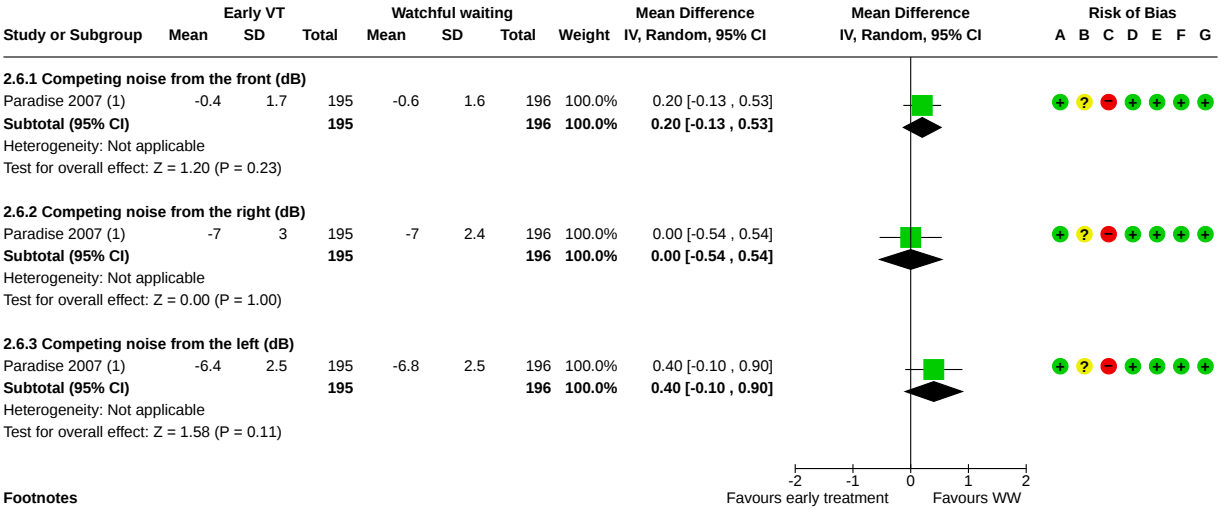
Test for subgroup differences: Not applicable



Footnotes
(1) Bilateral VT versus WW at 18 months, best ear at 4000Hz.
(2) At age 5. R and L ear data combined, with correction of variance. Assumed correlation coeff. of 0.5.
(3) Bilateral VT versus WW at 2 years. Maximum cases available.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 5: Mean final hearing threshold, randomised by child (long-term)

Analysis 2.6

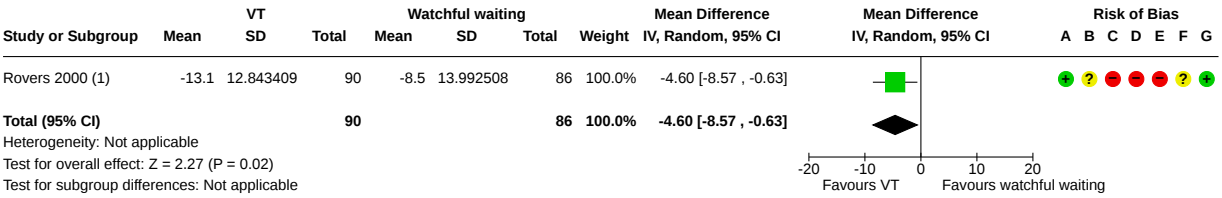


Footnotes
(1) Age 9 to 11.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 6: Hearing in noise test, randomised by child (long-term)

Analysis 2.7

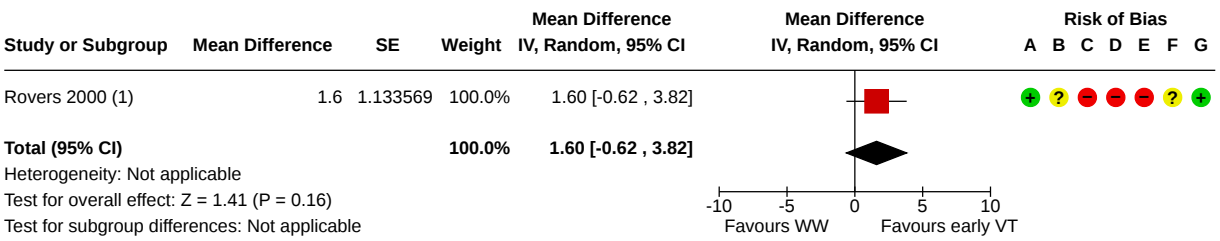


Footnotes
(1) Bilateral VT versus WW at 12 months. Better ear. Portable visual reinforcement audiology.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 7: Change in hearing threshold from baseline, randomised by child (medium-term)

Analysis 2.8



Footnotes
(1) At 12 months. Better ear. Adjusted for hearing level and age at randomization

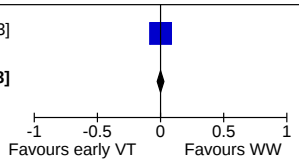
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 8: Adjusted mean difference in hearing improvement, randomised by child (medium term)

Analysis 2.9

Study or Subgroup	Early VT Events	Early VT Total	Watchful waiting Events	Watchful waiting Total	Weight	Risk Difference M-H, Random, 95% CI
Velepic 2011 (1)	0	59	0	102	100.0%	0.00 [-0.03 , 0.03]
Total (95% CI)		**59**		**102**	**100.0%**	**0.00 [-0.03 , 0.03]**
Total events:	0		0			

Heterogeneity: Not applicable
Test for overall effect: Z = 0.00 (P = 1.00)
Test for subgroup differences: Not applicable



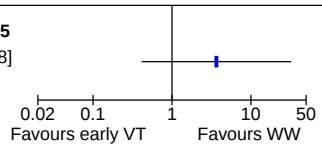
Footnotes
(1) Early VT + Ad versus WW + Ad, at least 6-months after surgery. Analysis by ears.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 9: Adverse event: persistent perforation, randomised by child (medium-term)

Analysis 2.10

Study or Subgroup	Early VT Events	Early VT Total	Watchful waiting Events	Watchful waiting Total	Risk Ratio M-H, Random, 95% CI
2.10.1 Adjusted for non-independence of within-individual measurements: ICC 0.5					
Paradise 2007 (1)	4	196	1	179	3.65 [0.41 , 32.38]



Footnotes
(1) At age 5. Analysis by ears. Each child contributed 2 data points, so average cluster size=2. DE=1.5

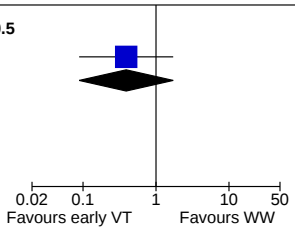
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 10: Adverse event: persistent perforation, randomised by child (long-term)

Analysis 2.11

Study or Subgroup	VT Events	VT Total	Watchful waiting Events	Watchful waiting Total	Weight	Risk Ratio M-H, Random, 95% CI
2.11.1 Adjusted for non-independence of within-individual measurements, assuming ICC of 0.5						
Velepic 2011 (1)	2	41	9	72	100.0%	0.39 [0.09 , 1.72]
Subtotal (95% CI)		**41**		**72**	**100.0%**	**0.39 [0.09 , 1.72]**
Total events:	2		9			

Heterogeneity: Not applicable
Test for overall effect: Z = 1.24 (P = 0.21)



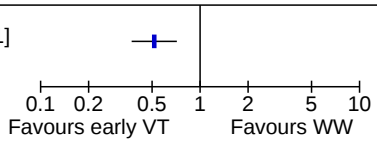
Footnotes
(1) At least 6 mo after surgery. Analysed by ear. Ave cluster size=1.85. DE=1.425

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 11: Presence/persistence of OME, randomised by child, measured by otoscopy (medium-term)

Analysis 2.12

Study or Subgroup	Early VT Events	Early VT Total	Watchful waiting Events	Watchful waiting Total	Risk Ratio M-H, Random, 95% CI
Maw 1999 (1)	29	80	52	74	0.52 [0.37 , 0.71]



Footnotes

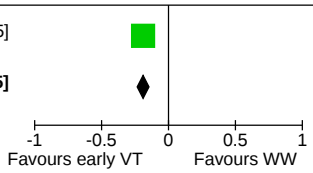
(1) Early VT versus WW. Effusion in best ear at 9 months.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 12: Presence/persistence of OME, randomised by child, measured by tympanometry (medium-term)

Analysis 2.13

Study or Subgroup	Early VT Mean	Early VT SD	Early VT Total	Watchful waiting Mean	Watchful waiting SD	Watchful waiting Total	Weight	Mean Difference IV, Random, 95% CI
Paradise 2007 (1)	0.29	0.2	159	0.48	0.2	157	100.0%	-0.19 [-0.23 , -0.15]
Total (95% CI)			**159**			**157**	**100.0%**	**-0.19 [-0.23 , -0.15]**

Heterogeneity: Not applicable
Test for overall effect: Z = 8.44 (P < 0.00001)
Test for subgroup differences: Not applicable



Footnotes

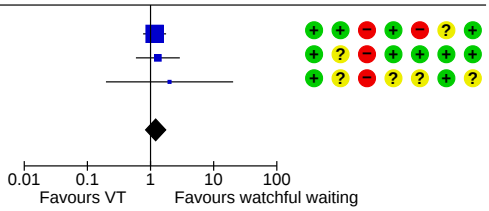
(1) Either uni-or bilateral effusion at age 3. Adjusted for laterality of effusion.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 13: Presence/persistence of OME, mean percentage of days, randomised by child (medium-term)

Analysis 2.14

Study or Subgroup	VT Events	VT Total	Watchful waiting Events	Watchful waiting Total	Weight	Risk Ratio M-H, Random, 95% CI
Maw 1999 (1)	33	83	24	70	76.5%	1.16 [0.76 , 1.76]
Paradise 2007 (2)	13	195	10	196	21.0%	1.31 [0.59 , 2.91]
Rach 1991 (3)	2	20	1	20	2.5%	2.00 [0.20 , 20.33]
Total (95% CI)		**298**		**286**	**100.0%**	**1.21 [0.84 , 1.74]**
Total events:	48		35			

Heterogeneity: Tau² = 0.00; Chi² = 0.26, df = 2 (P = 0.88); I² = 0%
Test for overall effect: Z = 1.00 (P = 0.32)
Test for subgroup differences: Not applicable



Footnotes

(1) Effusion in the better ear at 18 months by tympanometry.
(2) Effusion in either ear at age 9 to 11. Method of examination not reported.
(3) Bilateral flat tympanogram at age 7 to 8.

Risk of bias legend

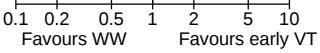
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 14: Presence/persistence of OME, randomised by child (long-term)

Analysis 2.15

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI	Odds Ratio IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	-0.01005	0.536779	100.0%	0.99 [0.35 , 2.83]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.99 [0.35 , 2.83]**		

Heterogeneity: Not applicable
Test for overall effect: Z = 0.02 (P = 0.99)
Test for subgroup differences: Not applicable



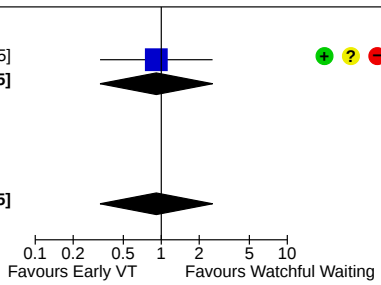
Footnotes
(1) WW versus early VT at age 7 to 8 years. Adjusted for age, gender, maternal education, housing, parity

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 15: Presence/persistence of OME, adjusted OR, randomised by child (long-term)

Analysis 2.16

Study or Subgroup	Early VT Events	Early VT Total	Watchful Waiting Events	Watchful Waiting Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
2.16.1 Adjusted for non-independence of within-individual measurements: ICC 0.5								
Paradise 2007 (1)	7	196	7	179	100.0%	0.91 [0.33 , 2.55]		+ ? − + + + +
Subtotal (95% CI)		**196**		**179**	**100.0%**	**0.91 [0.33 , 2.55]**		
Total events:	7		7					
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.17 (P = 0.86)								
Total (95% CI)		**196**		**179**	**100.0%**	**0.91 [0.33 , 2.55]**		
Total events:	7		7					
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.17 (P = 0.86)								
Test for subgroup differences: Not applicable								



Footnotes
(1) At age 5. Assessed using otomicrosopy. Ave cluster size=2. DE=1.5

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

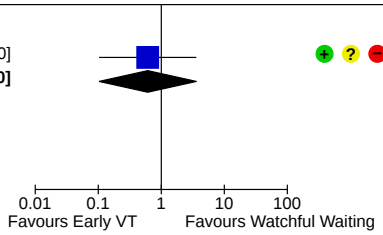
Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 16: Adverse event: tympanosclerosis (long term)

Analysis 2.17

Study or Subgroup	Early VT Events	Total	Watchful Waiting Events	Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
2.17.1 Adjusted for non-independence of within-individual measurements: ICC 0.5								
Paradise 2007 (1)	2	196	3	179	100.0%	0.61 [0.10 , 3.60]		
Subtotal (95% CI)		**196**		**179**	**100.0%**	**0.61 [0.10 , 3.60]**		
Total events:	2		3					

Heterogeneity: Not applicable
Test for overall effect: Z = 0.55 (P = 0.58)

Test for subgroup differences: Not applicable



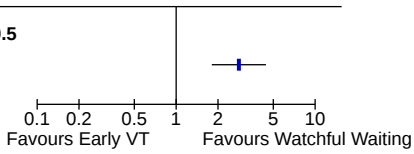
Footnotes
(1) assessed using otomicrosopy. Ave cluster size=2. DE=1.5

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 17: Adverse event: fibrosis (long term)

Analysis 2.18

Study or Subgroup	Early VT Events	Total	Watchful Waiting Events	Total	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
2.18.1 Adjusted for non-independence of within-individual measurements. Assumed ICC 0.5						
Paradise 2007 (1)	65	196	21	179	2.83 [1.81 , 4.43]	



Footnotes
(1) Age 5 years. assessed using otomicrosopy. Ave cluster size=2. DE=1.5

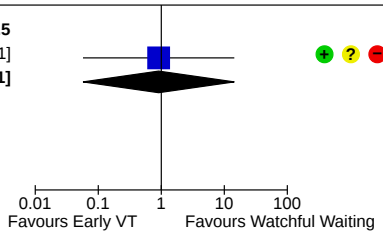
Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 18: Adverse event: segmental atrophy (long term)

Analysis 2.19

Study or Subgroup	Early VT Events	Total	Watchful Waiting Events	Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
2.19.1 Adjusted for non-independence of within-individual measurements. ICC assumed 0.5								
Paradise 2007 (1)	1	196	1	178	100.0%	0.91 [0.06 , 14.41]		
Subtotal (95% CI)		**196**		**178**	**100.0%**	**0.91 [0.06 , 14.41]**		
Total events:	1		1					

Heterogeneity: Not applicable
Test for overall effect: Z = 0.07 (P = 0.95)

Test for subgroup differences: Not applicable



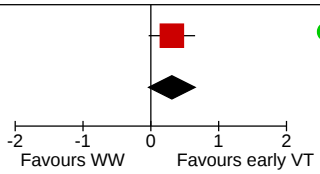
Footnotes
(1) Reported by ears. Assessed using otomicrosopy. Ave cluster size=2. DE=1.5

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 19: Adverse event: retraction pocket with other abnormality (long term)

Analysis 2.20

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.31	0.174709	100.0%	0.31 [-0.03 , 0.65]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.31 [-0.03 , 0.65]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 1.77 (P = 0.08)
Test for subgroup differences: Not applicable

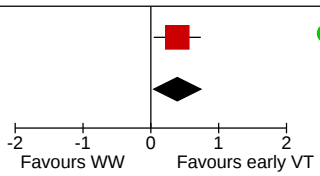
Footnotes
(1) Reynell test at 9 months. Mean difference between groups for deficit from chronological age

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 20: Receptive language development, Reynell test, randomised by child (medium-term)

Analysis 2.21

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.39	0.176818	100.0%	0.39 [0.04 , 0.74]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.39 [0.04 , 0.74]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 2.21 (P = 0.03)
Test for subgroup differences: Not applicable

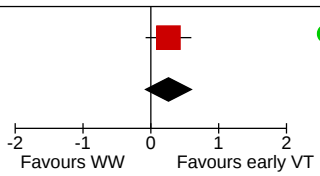
Footnotes
(1) Reynell test at 9 months, adjusted for age, sex and hearing at randomisation.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 21: Receptive language development, Reynell test, adj MD (medium-term)

Analysis 2.22

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.26	0.172073	100.0%	0.26 [-0.08 , 0.60]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.26 [-0.08 , 0.60]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 1.51 (P = 0.13)
Test for subgroup differences: Not applicable

Footnotes
(1) Reynell test (standardised score) at 18 months.

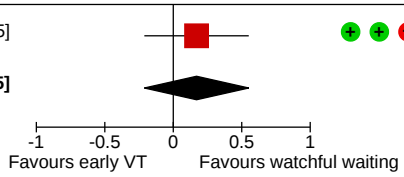
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 22: Receptive language, Reynell test, randomised by child (long-term)

Analysis 2.23

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI
Maw 1999 (1)	0.17	0.1945	100.0%	0.17 [-0.21 , 0.55]
Total (95% CI)			**100.0%**	**0.17 [-0.21 , 0.55]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.87 (P = 0.38)
Test for subgroup differences: Not applicable



Risk of Bias: A B C D E F G

Footnotes
(1) Reynell test at 18 months. Adjusted for age, sex, hearing

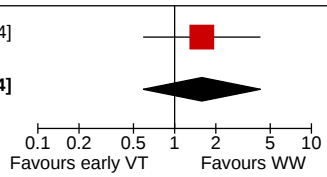
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 23: Receptive language: Reynell test, long-term, adjusted MD

Analysis 2.24

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
Maw 1999 (1)	0.457425	0.503721	100.0%	1.58 [0.59 , 4.24]
Total (95% CI)			**100.0%**	**1.58 [0.59 , 4.24]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.91 (P = 0.36)
Test for subgroup differences: Not applicable



Risk of Bias: A B C D E F G

Footnotes
(1) WOLD at age 7 to 8. WW versus early surgery. Adjusted for age, gender, maternal education, housing and mother’s parity.

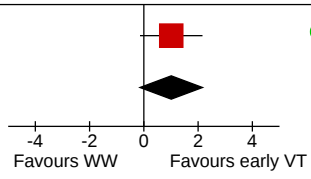
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 24: Receptive language: WOLD adjusted OR (long-term)

Analysis 2.25

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI
Rovers 2000 (1)	1.01	0.587764	100.0%	1.01 [-0.14 , 2.16]
Total (95% CI)			**100.0%**	**1.01 [-0.14 , 2.16]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.72 (P = 0.09)
Test for subgroup differences: Not applicable



Risk of Bias: A B C D E F G

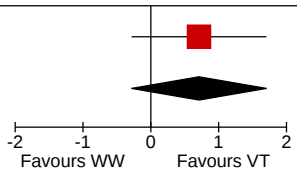
Footnotes
(1) VT versus WW at 12 months.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 25: Receptive language, mean difference (months) in improvement in Reynell test score (equivalent age -real age): medium-term

Analysis 2.26

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Rovers 2000 (1)	0.71	0.506876	100.0%	0.71 [-0.28 , 1.70]		+ ? − − − ? +
Total (95% CI)			**100.0%**	**0.71 [-0.28 , 1.70]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 1.40 (P = 0.16)
Test for subgroup differences: Not applicable

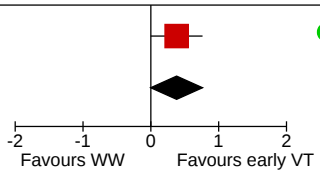
Footnotes
(1) At 12 months. Adjusted for IQ, baseline language development and maternal education.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 26: Receptive language, adjusted mean difference (months) in improvement in Reynell test score (equivalent age - real age): medium-term

Analysis 2.27

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.38	0.194956	100.0%	0.38 [-0.00 , 0.76]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.38 [-0.00 , 0.76]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 1.95 (P = 0.05)
Test for subgroup differences: Not applicable

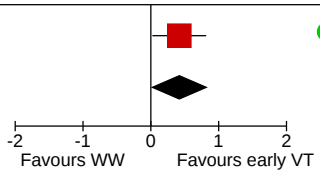
Footnotes
(1) Reynell test, standardised score at 9 months.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 27: Expressive language development: Reynell test (medium-term)

Analysis 2.28

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.42	0.202336	100.0%	0.42 [0.02 , 0.82]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.42 [0.02 , 0.82]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 2.08 (P = 0.04)
Test for subgroup differences: Not applicable

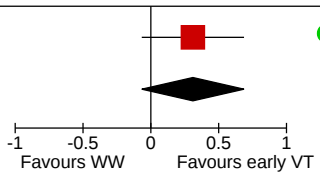
Footnotes
(1) Reynell test at 9 months, adjusted for age, sex and hearing at randomisation.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 28: Expressive language development: Reynell test, medium-term, adjusted MD

Analysis 2.29

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.31	0.192317	100.0%	0.31 [-0.07 , 0.69]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.31 [-0.07 , 0.69]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 1.61 (P = 0.11)
Test for subgroup differences: Not applicable

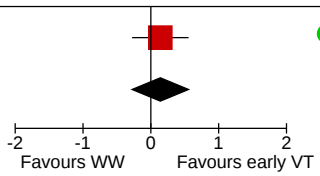
Footnotes
(1) Reynell test at 18 months.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 29: Expressive language development: Reynell test (long-term)

Analysis 2.30

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.14	0.212181	100.0%	0.14 [-0.28 , 0.56]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.14 [-0.28 , 0.56]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 0.66 (P = 0.51)
Test for subgroup differences: Not applicable

Footnotes
(1) Reynell test at 18 months, adjusted for age, sex and hearing at randomisation.

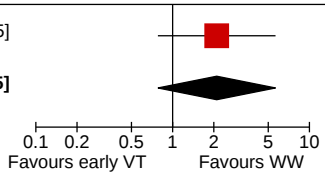
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 30: Expressive language development: Reynell test, long-term, adjusted MD

Analysis 2.31

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.741937	0.505141	100.0%	2.10 [0.78 , 5.65]	+ + − + − ? +
Total (95% CI)			**100.0%**	**2.10 [0.78 , 5.65]**	

Heterogeneity: Not applicable
Test for overall effect: Z = 1.47 (P = 0.14)
Test for subgroup differences: Not applicable



Footnotes
(1) WOLD at age 7 to 8. WW versus early surgery. Adjusted for age, gender, maternal education, housing and mother’s parity.

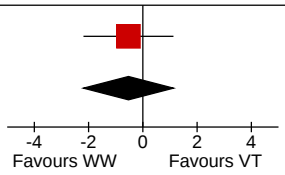
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 31: Expressive language: WOLD adjusted OR (long-term)

Analysis 2.32

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Rovers 2000 (1)	-0.53	0.846058	100.0%	-0.53 [-2.19 , 1.13]	+ ? − − − ? +
Total (95% CI)			**100.0%**	**-0.53 [-2.19 , 1.13]**	

Heterogeneity: Not applicable
Test for overall effect: Z = 0.63 (P = 0.53)
Test for subgroup differences: Not applicable



Footnotes
(1) VT versus WW at 12 months.

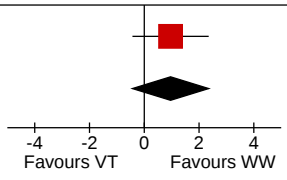
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 32: Expressive language, mean difference (months) in improvement in Schlichting test score (equivalent age -real age): medium-term

Analysis 2.33

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Rovers 2000 (1)	0.96	0.71	100.0%	0.96 [-0.43 , 2.35]	+ ? − − − ? +
Total (95% CI)			**100.0%**	**0.96 [-0.43 , 2.35]**	

Heterogeneity: Not applicable
Test for overall effect: Z = 1.35 (P = 0.18)
Test for subgroup differences: Not applicable



Footnotes
(1) At 12 months. Adjusted for IQ, baseline language development and maternal education.

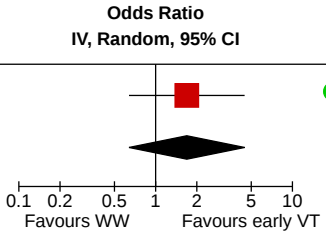
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

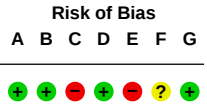
Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 33: Expressive language, adjusted mean difference (months) in improvement in Schlichting test score (equivalent age - real age): medium-term

Analysis 2.34

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
Maw 1999 (1)	0.524729	0.496415	100.0%	1.69 [0.64 , 4.47]
Total (95% CI)			**100.0%**	**1.69 [0.64 , 4.47]**

Heterogeneity: Not applicable
Test for overall effect: Z = 1.06 (P = 0.29)
Test for subgroup differences: Not applicable





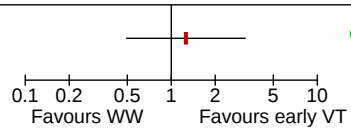
Footnotes
(1) CN/Rep at age 7-8. WW versus early surgery. Adj for age, gender, mat edu, housing, mother’s parity.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 34: Non-word repetition total score, adjusted OR (long-term)

Analysis 2.35

Study or Subgroup	log[Odds Ratio]	SE	Odds Ratio IV, Random, 95% CI
Maw 1999 (1)	0.231112	0.481088	1.26 [0.49 , 3.23]



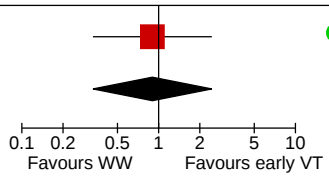
Footnotes
(1) Age 7 to 8. OR for WW versus early VT. Adj for age, gender, maternal education, housing and parity.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 35: Reading , WORD test, adjusted OR (long-term)

Analysis 2.36

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI	Odds Ratio IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	-0.105361	0.510382	100.0%	0.90 [0.33 , 2.45]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.90 [0.33 , 2.45]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 0.21 (P = 0.84)
Test for subgroup differences: Not applicable

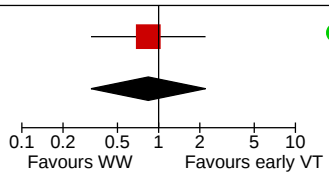
Footnotes
(1) Age 7 to 8. OR for WW versus early VT. Adj for age, gender, maternal education, housing and parity.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 36: Spelling, ALSPAC test, adjusted OR (long-term)

Analysis 2.37

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI	Odds Ratio IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	-0.174353	0.491818	100.0%	0.84 [0.32 , 2.20]		+ + − + − ? +
Total (95% CI)			**100.0%**	**0.84 [0.32 , 2.20]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 0.35 (P = 0.72)
Test for subgroup differences: Not applicable

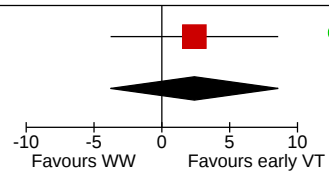
Footnotes
(1) Age 7 to 8. OR for WW versus early VT. Adj for age, gender, maternal education, housing and parity.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 37: Phoneme deletion, adjusted OR (long-term)

Analysis 2.38

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	2.4	3.151752	100.0%	2.40 [-3.78 , 8.58]		+ + − + − ? +
Total (95% CI)			**100.0%**	**2.40 [-3.78 , 8.58]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 0.76 (P = 0.45)
Test for subgroup differences: Not applicable

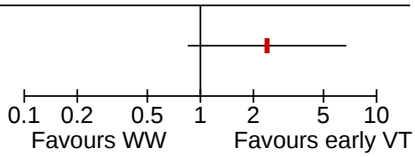
Footnotes
(1) Griffiths practical reasoning subscale at 9 months.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 38: Cognitive development: Griffiths practical reasoning (medium-term)

Analysis 2.39

Study or Subgroup	log[Odds Ratio]	SE	Odds Ratio IV, Random, 95% CI	Odds Ratio IV, Random, 95% CI
Maw 1999 (1)	0.871293	0.528974	2.39 [0.85 , 6.74]	



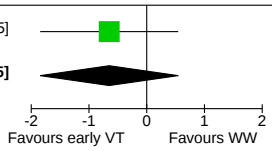
Footnotes

(1) Total IQ at age 7-8. Adj for age, gender, mat edu, housing, mother’s parity.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 39: Cognitive development: IQ (WISC-III UK short form) adjusted OR (long term)

Analysis 2.40

Study or Subgroup	Early VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	8.21	3.22	84	8.86	4.05	66	100.0%	-0.65 [-1.85 , 0.55]		+ + − + − ? +
Total (95% CI)			**84**			**66**	**100.0%**	**-0.65 [-1.85 , 0.55]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 1.07 (P = 0.29)
Test for subgroup differences: Not applicable

Footnotes

(1) Bilateral VT versus WW at 9 months.

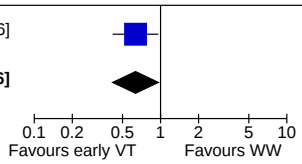
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 40: Behaviour, Richman score (medium-term)

Analysis 2.41

Study or Subgroup	Early VT Events	Total	Watchful waiting Events	Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	25	84	31	66	100.0%	0.63 [0.42 , 0.96]		+ + − + − ? +
Total (95% CI)		**84**		**66**	**100.0%**	**0.63 [0.42 , 0.96]**		
Total events:	25		31					



Heterogeneity: Not applicable
Test for overall effect: Z = 2.15 (P = 0.03)
Test for subgroup differences: Not applicable

Footnotes

(1) Bilateral VT versus WW at 9 months. Dependent variable problem present (>/=10) or absent.

Risk of bias legend

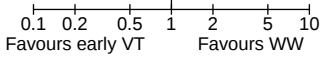
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 41: Behaviour, Richman score, dichotomised (medium-term)

Analysis 2.42

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI	Odds Ratio IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.14842	0.73514	100.0%	1.16 [0.27 , 4.90]		
Total (95% CI)			**100.0%**	**1.16 [0.27 , 4.90]**		

Heterogeneity: Not applicable
Test for overall effect: Z = 0.20 (P = 0.84)
Test for subgroup differences: Not applicable



Footnotes
(1) At 9 mo. Adj for baseline hearing, age, duration of HL, 7-mo old hearing screening and current HL.

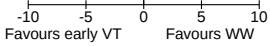
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 42: Behaviour, Richman score, adjusted OR (medium-term)

Analysis 2.43

Study or Subgroup	VT Mean	VT SD	VT Total	Watchful waiting Mean	Watchful waiting SD	Watchful waiting Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	7.9	3.01	67	7	3.5	56	100.0%	0.90 [-0.27 , 2.07]		
Total (95% CI)			**67**			**56**	**100.0%**	**0.90 [-0.27 , 2.07]**		

Heterogeneity: Not applicable
Test for overall effect: Z = 1.51 (P = 0.13)
Test for subgroup differences: Not applicable



Footnotes
(1) Bilateral VT versus WW at 18 months

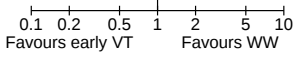
Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 43: Behaviour, Richman score (long-term)

Analysis 2.44

Study or Subgroup	Early VT Events	Early VT Total	Watchful waiting Events	Watchful waiting Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	16	67	11	56	100.0%	1.22 [0.62 , 2.40]		
Total (95% CI)		**67**		**56**	**100.0%**	**1.22 [0.62 , 2.40]**		
Total events:	16		11					

Heterogeneity: Not applicable
Test for overall effect: Z = 0.56 (P = 0.57)
Test for subgroup differences: Not applicable



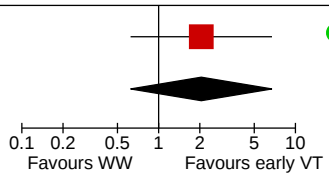
Footnotes
(1) Bilateral VT versus WW at 18 months. Dependent variable problem present (>/=10) or absent.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 44: Behaviour, Richman score, dichotomised (long-term)

Analysis 2.45

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI	Odds Ratio IV, Random, 95% CI	Risk of Bias A B C D E F G
Maw 1999 (1)	0.71784	0.607191	100.0%	2.05 [0.62 , 6.74]		+ + − + − ? +
Total (95% CI)			**100.0%**	**2.05 [0.62 , 6.74]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 1.18 (P = 0.24)
Test for subgroup differences: Not applicable

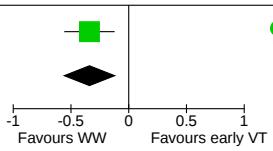
Footnotes
(1) SDQ (teacher, total) at age 7-8, adj for age, gender, mat edu, housing, mother’s parity.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 45: Behaviour: SDQ teacher report, total, adjusted OR (long-term)

Analysis 2.46

Study or Subgroup	VT Mean	VT SD	VT Total	Watchful waiting Mean	Watchful waiting SD	Watchful waiting Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Rovers 2000 (1)	5.88	0.799	84	6.22	0.622	81	100.0%	-0.34 [-0.56 , -0.12]		+ ? − − − ? +
Total (95% CI)			**84**			**81**	**100.0%**	**-0.34 [-0.56 , -0.12]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 3.06 (P = 0.002)
Test for subgroup differences: Not applicable

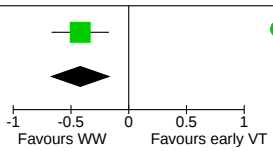
Footnotes
(1) At 12 months. Combined means across five domains, with correction of variance. Assumed correlation coeff. of 0.5. Higher = better.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 46: Parent-child interaction: Erickson child scale (medium-term)

Analysis 2.47

Study or Subgroup	Early VT Mean	Early VT SD	Early VT Total	Watchful waiting Mean	Watchful waiting SD	Watchful waiting Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Rovers 2000 (1)	5.3	0.875	84	5.72	0.749	81	100.0%	-0.42 [-0.67 , -0.17]		+ ? − − − ? +
Total (95% CI)			**84**			**81**	**100.0%**	**-0.42 [-0.67 , -0.17]**		



Heterogeneity: Not applicable
Test for overall effect: Z = 3.32 (P = 0.0009)
Test for subgroup differences: Not applicable

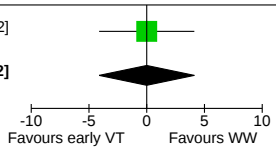
Footnotes
(1) At 12 months. Combined means across five domains, with correction of variance. Assumed correlation coeff. of 0.5. Higher = better.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 47: Parent-child interaction: Erickson parent scale (medium-term)

Analysis 2.48

Study or Subgroup	Early VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Paradise 2007 (1)	66	19	194	66	22	189	100.0%	0.00 [-4.12 , 4.12]		+ ? − + + + +
Total (95% CI)			**194**			**189**	**100.0%**	**0.00 [-4.12 , 4.12]**		



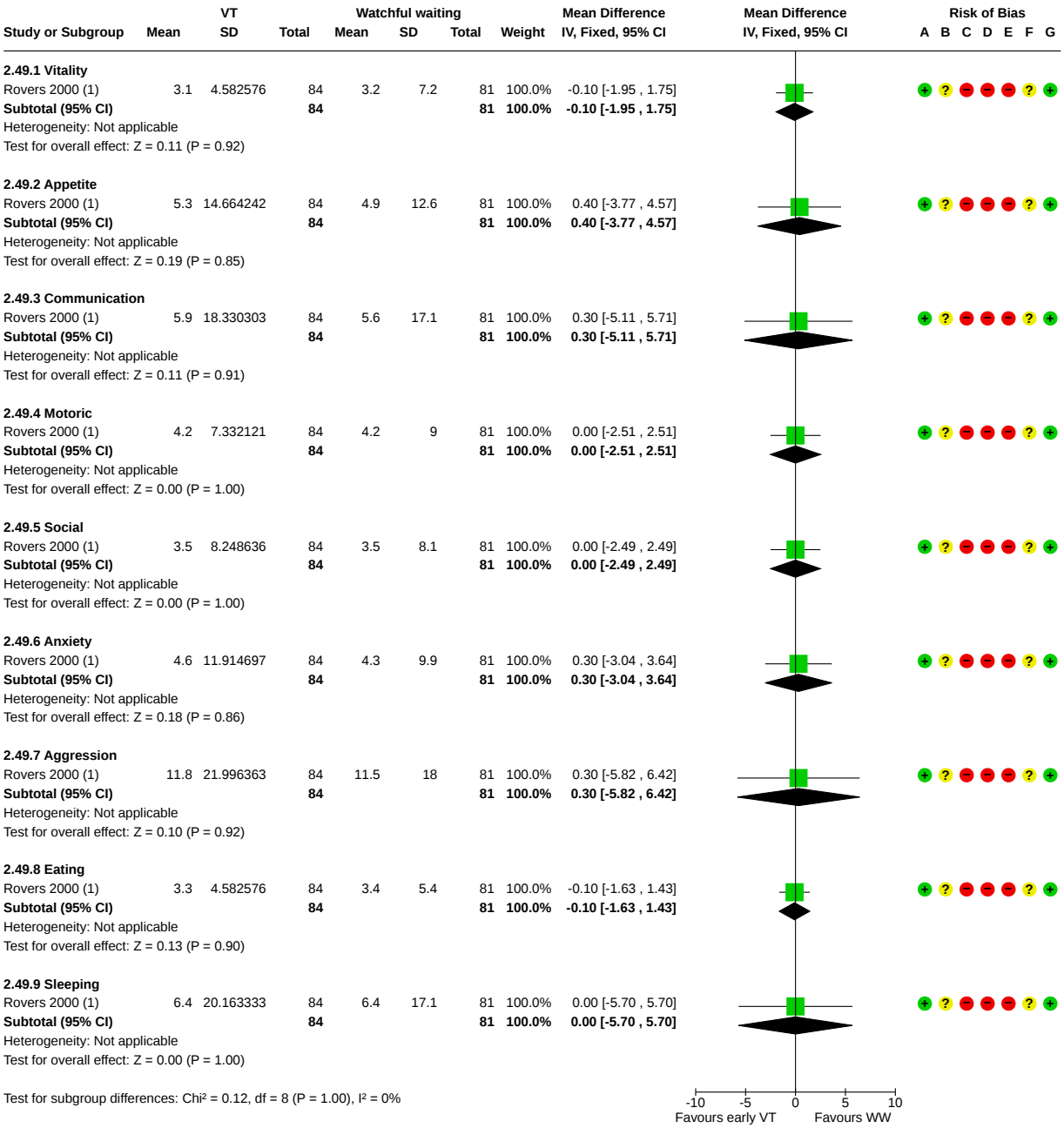
Heterogeneity: Not applicable
Test for overall effect: Z = 0.00 (P = 1.00)
Test for subgroup differences: Not applicable

Footnotes
(1) Total stress score at age 6.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 48: Parental stress, Parental Stress Index, short form (long-term)

Analysis 2.49



Footnotes

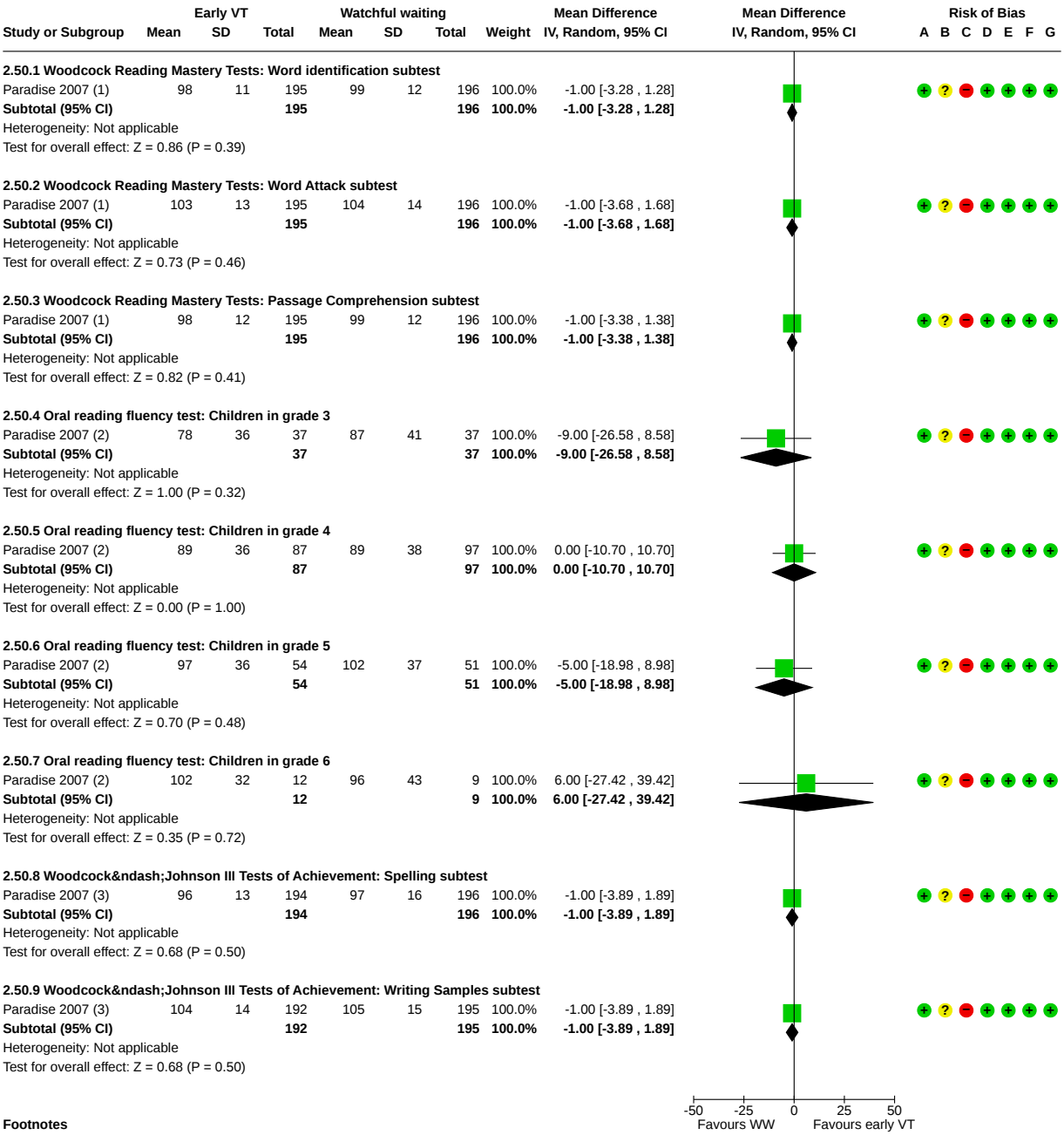
(1) At 12 months. Higher score = more complaints.

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 49: Generic health-related quality of life: TAIQOL (medium-term)

Analysis 2.50



Footnotes

(1) Age 9 to 11. The normative mean standard score is 100±15. Higher scores indicate more favorable results.
(2) Age 9 to 11. Higher scores indicate more favorable results.
(3) Age 9 to 11. The normative mean standard score on both subtests is 100±15. Higher scores indicate more favorable results.

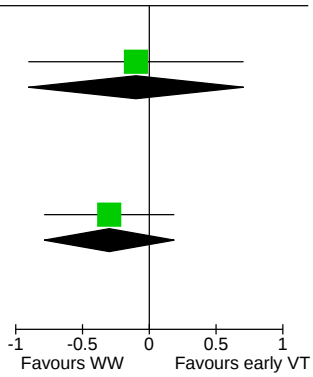
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 50: Literacy (long-term)

Analysis 2.51

Study or Subgroup	Early VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.51.1 Comprehensive Test of Phonological Processing: Elision subtest								
Paradise 2007 (1)	8.6	4.9	195	8.7	3	196	100.0%	-0.10 [-0.91 , 0.71]
Subtotal (95% CI)			**195**			**196**	**100.0%**	**-0.10 [-0.91 , 0.71]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.24 (P = 0.81)								
2.51.2 Comprehensive Test of Phonological Processing: Rapid Letter Naming subtest								
Paradise 2007 (1)	9.3	2.5	193	9.6	2.4	196	100.0%	-0.30 [-0.79 , 0.19]
Subtotal (95% CI)			**193**			**196**	**100.0%**	**-0.30 [-0.79 , 0.19]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.21 (P = 0.23)								



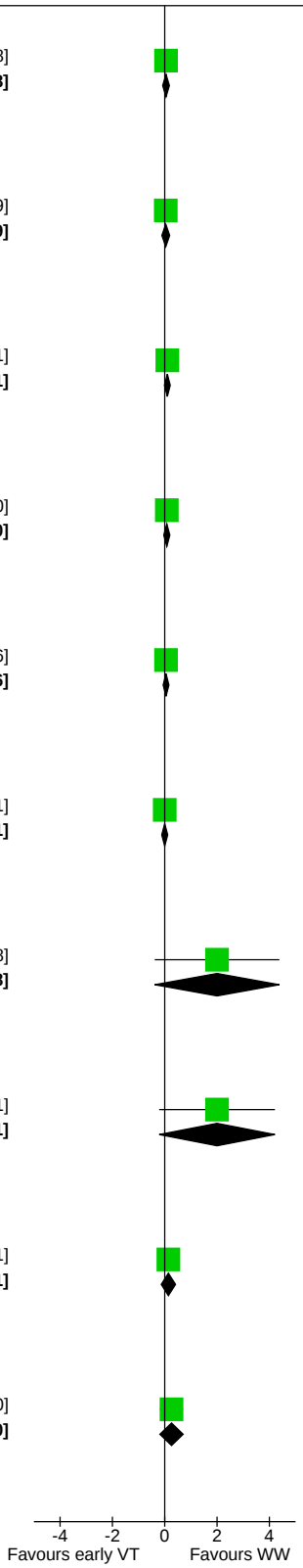
Footnotes

(1) Age 9 to 11. Higher scores indicate more favourable results.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 51: Phonological awareness (long-term)

Analysis 2.52

Study or Subgroup	Early VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.52.1 Disruptive Behavior Disorders Rating Scale: Inattention factor: Parent's rating								
Paradise 2007 (1)	0.7	0.63	194	0.65	0.66	196	100.0%	0.05 [-0.08 , 0.18]
Subtotal (95% CI)			**194**			**196**	**100.0%**	**0.05 [-0.08 , 0.18]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.77 (P = 0.44)								
2.52.2 Disruptive Behavior Disorders Rating Scale: Inattention factor: Teacher's rating								
Paradise 2007 (1)	0.71	0.74	190	0.67	0.75	192	100.0%	0.04 [-0.11 , 0.19]
Subtotal (95% CI)			**190**			**192**	**100.0%**	**0.04 [-0.11 , 0.19]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.52 (P = 0.60)								
2.52.3 Disruptive Behavior Disorders Rating Scale: Impulsivity and overactivity factor: Parent's rating								
Paradise 2007 (1)	0.67	0.57	194	0.57	0.54	196	100.0%	0.10 [-0.01 , 0.21]
Subtotal (95% CI)			**194**			**196**	**100.0%**	**0.10 [-0.01 , 0.21]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.78 (P = 0.08)								
2.52.4 Disruptive Behavior Disorders Rating Scale: Impulsivity and overactivity factor: Teacher's rating								
Paradise 2007 (1)	0.48	0.63	190	0.4	0.52	192	100.0%	0.08 [-0.04 , 0.20]
Subtotal (95% CI)			**190**			**192**	**100.0%**	**0.08 [-0.04 , 0.20]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.35 (P = 0.18)								
2.52.5 Disruptive Behavior Disorders Rating Scale: Oppositional defiant factor: Parent's rating								
Paradise 2007 (1)	0.57	0.58	194	0.52	0.53	196	100.0%	0.05 [-0.06 , 0.16]
Subtotal (95% CI)			**194**			**196**	**100.0%**	**0.05 [-0.06 , 0.16]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.89 (P = 0.37)								
2.52.6 Disruptive Behavior Disorders Rating Scale: Oppositional defiant factor: Teacher's rating								
Paradise 2007 (1)	0.33	0.56	190	0.33	0.58	192	100.0%	0.00 [-0.11 , 0.11]
Subtotal (95% CI)			**190**			**192**	**100.0%**	**0.00 [-0.11 , 0.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								
2.52.7 Child Behavior Checklist: Total Problems score, parent’s rating								
Paradise 2007 (2)	51	12	194	49	12	196	100.0%	2.00 [-0.38 , 4.38]
Subtotal (95% CI)			**194**			**196**	**100.0%**	**2.00 [-0.38 , 4.38]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.65 (P = 0.10)								
2.52.8 Child Behavior Checklist: Total Problems score, teacher’s rating								
Paradise 2007 (2)	52	11	189	50	11	191	100.0%	2.00 [-0.21 , 4.21]
Subtotal (95% CI)			**189**			**191**	**100.0%**	**2.00 [-0.21 , 4.21]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.77 (P = 0.08)								
2.52.9 Impairment Rating Scales: Overall functioning, parent’s rating								
Paradise 2007 (3)	0.82	1.42	194	0.68	1.33	196	100.0%	0.14 [-0.13 , 0.41]
Subtotal (95% CI)			**194**			**196**	**100.0%**	**0.14 [-0.13 , 0.41]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.00 (P = 0.32)								
2.52.10 Impairment Rating Scales: Overall functioning, teacher’s rating								
Paradise 2007 (3)	2.04	2.24	190	1.78	2.19	192	100.0%	0.26 [-0.18 , 0.70]
Subtotal (95% CI)			**190**			**192**	**100.0%**	**0.26 [-0.18 , 0.70]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.15 (P = 0.25)								



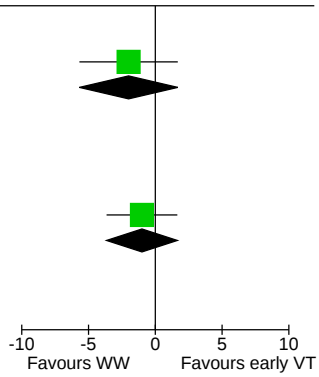
Footnotes

(1) Age 9 to 11. Higher scores indicate less favourable results.
(2) Age 9 to 11. Higher scores indicate less favorable results.
(3) Age 9 to 11. A score of 3 or higher is considered to be indicative of clinically meaningful impairment.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 52: Attention, impulsivity, and psychosocial function, long-term (1): disruptive behaviour disorders, child behaviour and impairment rating

Analysis 2.53

Study or Subgroup	Early VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.53.1 Attention, impulsivity, and psychosocial function: Social Skills Rating System: parent version								
Paradise 2007 (1)	96	19	194	98	18	194	100.0%	-2.00 [-5.68 , 1.68]
Subtotal (95% CI)			**194**			**194**	**100.0%**	**-2.00 [-5.68 , 1.68]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.06 (P = 0.29)								
2.53.2 Attention, impulsivity, and psychosocial function: Social Skills Rating System: teacher version								
Paradise 2007 (1)	98	13	184	99	13	186	100.0%	-1.00 [-3.65 , 1.65]
Subtotal (95% CI)			**184**			**186**	**100.0%**	**-1.00 [-3.65 , 1.65]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.74 (P = 0.46)								
Test for subgroup differences: Chi² = 0.19, df = 1 (P = 0.67), I² = 0%								



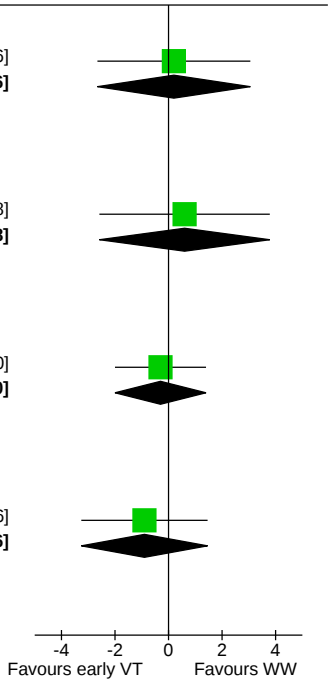
Footnotes

(1) At age 9 to 11. The normative mean standard score is 100±15. Higher scores indicate more favorable results.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 53: Attention, impulsivity, and psychosocial function, long-term (2): social skills

Analysis 2.54

Study or Subgroup	Early VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.54.1 Visual Continuous Performance Test: Inattention								
Paradise 2007 (1)	9.7	18.5	195	9.5	8.5	196	100.0%	0.20 [-2.66 , 3.06]
Subtotal (95% CI)			**195**			**196**	**100.0%**	**0.20 [-2.66 , 3.06]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.14 (P = 0.89)								
2.54.2 Visual Continuous Performance Test: Impulsivity								
Paradise 2007 (1)	8.8	16.5	195	8.2	15.6	196	100.0%	0.60 [-2.58 , 3.78]
Subtotal (95% CI)			**195**			**196**	**100.0%**	**0.60 [-2.58 , 3.78]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.37 (P = 0.71)								
2.54.3 Auditory Continuous Performance Test: Inattention								
Paradise 2007 (2)	11.1	7.2	155	11.4	8	153	100.0%	-0.30 [-2.00 , 1.40]
Subtotal (95% CI)			**155**			**153**	**100.0%**	**-0.30 [-2.00 , 1.40]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.35 (P = 0.73)								
2.54.4 Auditory Continuous Performance Test: Impulsivity								
Paradise 2007 (2)	3.3	8.7	154	4.2	12.1	153	100.0%	-0.90 [-3.26 , 1.46]
Subtotal (95% CI)			**154**			**153**	**100.0%**	**-0.90 [-3.26 , 1.46]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.75 (P = 0.45)								



Footnotes

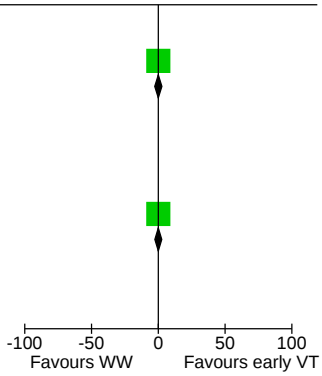
(1) At age 9 to 11. Higher scores indicate less favorable results.

(2) At age 9 to 11. Higher scores indicate less favorable results.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 54: Attention, impulsivity, and psychosocial function, long-term: Visual and auditory continuous performance

Analysis 2.55

Study or Subgroup	Early VT Mean	SD	Total	Watchful waiting Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
2.55.1 Wechsler Abbreviated Scale of Intelligence								
Paradise 2007 (1)	96	13	195	96	14	196	100.0%	0.00 [-2.68 , 2.68]
Subtotal (95% CI)			**195**			**196**	**100.0%**	**0.00 [-2.68 , 2.68]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								
2.55.2 Calculation subtest of the Woodcock–Johnson III Tests of Achievement								
Paradise 2007 (2)	99	13	194	99	13	195	100.0%	0.00 [-2.58 , 2.58]
Subtotal (95% CI)			**194**			**195**	**100.0%**	**0.00 [-2.58 , 2.58]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.00 (P = 1.00)								



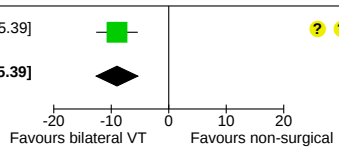
Footnotes

(1) At age 9 to 11. The normative mean score is 100±15. Higher scores indicate more favorable results.
(2) At age 9 to 11. The normative mean score is 100±15. Higher scores indicate more favorable results.

Comparison 2: Early VT versus watchful waiting (treatment later if required) , Outcome 55: Intelligence and academic achievement (long-term)

Analysis 3.1

Study or Subgroup	Bilateral VT Mean	SD	Total	Non-surgical treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Bernard 1991 (1)	11.5	10.3	60	20.5	10.3	65	100.0%	-9.00 [-12.61 , -5.39]	? ? − − + ? −
Total (95% CI)			**60**			**65**	**100.0%**	**-9.00 [-12.61 , -5.39]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 4.88 (P < 0.00001)									
Test for subgroup differences: Not applicable									



Footnotes

(1) At 2 months. The non-surgical treatment was antibiotic (sulfisoxazole)

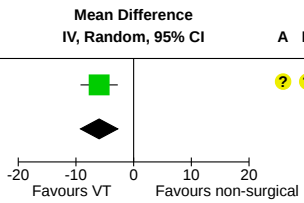
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 3: VT versus non-surgical treatment, Outcome 1: Mean final hearing threshold (short-term)

Analysis 3.2

Study or Subgroup	VT Mean	SD	Total	Non-surgical treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Bernard 1991 (1)	11.83	9.2	60	17.81	9.2	65	100.0%	-5.98 [-9.21 , -2.75]	? ? − − + ? −
Total (95% CI)			**60**			**65**	**100.0%**	**-5.98 [-9.21 , -2.75]**	
Heterogeneity: Not applicable									
Test for overall effect: Z = 3.63 (P = 0.0003)									
Test for subgroup differences: Not applicable									



Footnotes

(1) At 4 months. The non-surgical treatment was antibiotic (sulfisoxazole)

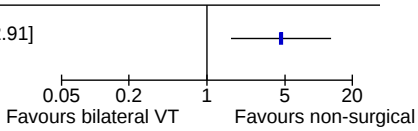
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 3: VT versus non-surgical treatment, Outcome 2: Mean final hearing threshold (medium-term)

Analysis 3.3

Study or Subgroup	Bilateral VT Events	Total	non-surgical treatment Events	Total	Risk Ratio M-H, Random, 95% CI
Bernard 1991 (1)	17	60	4	65	4.60 [1.64 , 12.91]



Footnotes

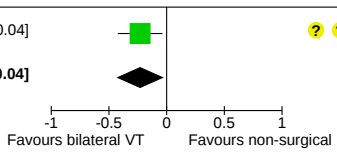
(1) VT versus sulfisoxazole at 18 months.

Comparison 3: VT versus non-surgical treatment, Outcome 3: Adverse event: myringosclerosis (long-term)

Analysis 3.4

Study or Subgroup	Bilateral VT Mean	SD	Total	Non-surgical treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Bernard 1991 (1)	0.33	0.55	60	0.56	0.55	65	100.0%	-0.23 [-0.42 , -0.04]	? ? − − + ? −
Total (95% CI)			**60**			**65**	**100.0%**	**-0.23 [-0.42 , -0.04]**	

Heterogeneity: Not applicable
Test for overall effect: Z = 2.34 (P = 0.02)
Test for subgroup differences: Not applicable



Footnotes

(1) At 6 to 12 months

Risk of bias legend

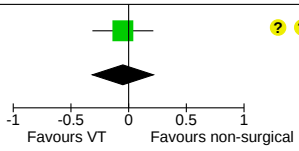
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 3: VT versus non-surgical treatment, Outcome 4: Number of doctor-diagnosed AOM episodes (medium-term)

Analysis 3.5

Study or Subgroup	VT Mean	SD	Total	Non-surgical treatment Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Bernard 1991 (1)	0.37	1	60	0.42	0.31	65	100.0%	-0.05 [-0.31 , 0.21]	? ? − − + ? −
Total (95% CI)			**60**			**65**	**100.0%**	**-0.05 [-0.31 , 0.21]**	

Heterogeneity: Not applicable
Test for overall effect: Z = 0.37 (P = 0.71)
Test for subgroup differences: Not applicable



Footnotes

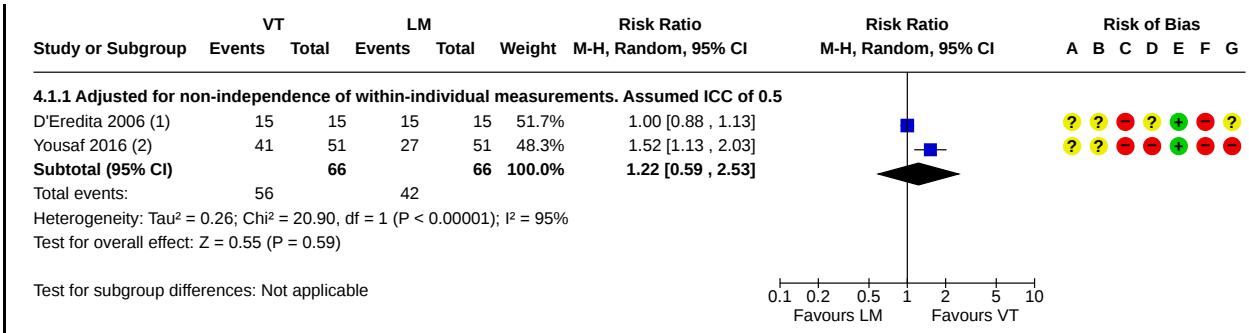
(1) At 12 to 18 months

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 3: VT versus non-surgical treatment, Outcome 5: Number of doctor-diagnosed episodes of AOM (long-term)

Analysis 4.1



Footnotes

(1) Bilateral VT versus bilateral laser myringotomy at one year follow-up.

(2) At 6 months. Unilateral or bilateral treatment in each group. Reported by ear. Average cluster size=1.66. DE=1.33

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

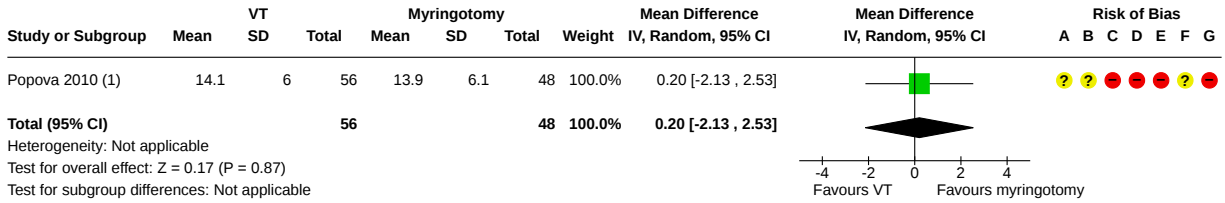
(E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias)

(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 1: Hearing returned to normal: VT versus laser myringotomy (medium-term)

Analysis 4.2



Footnotes

(1) VT + AD versus Myringotomy + Ad. Randomised by child, reported by ear at 1 month. Ave cluster size=2. Assumed ICC 0.5. DE=1.5

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

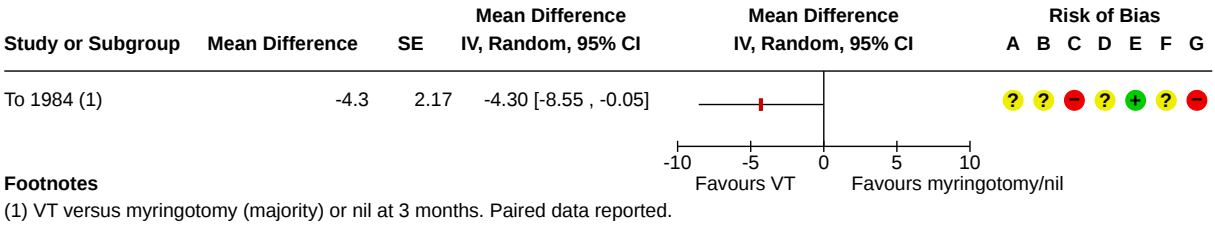
(E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias)

(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 2: Mean final hearing threshold, randomised by child (short-term). Adjusted for non-independence of within-individual measurements. Assumed ICC of 0.5

Analysis 4.3



Footnotes

(1) VT versus myringotomy (majority) or nil at 3 months. Paired data reported.

Risk of bias legend

(A) Random sequence generation (selection bias)

(B) Allocation concealment (selection bias)

(C) Blinding of participants and personnel (performance bias)

(D) Blinding of outcome assessment (detection bias)

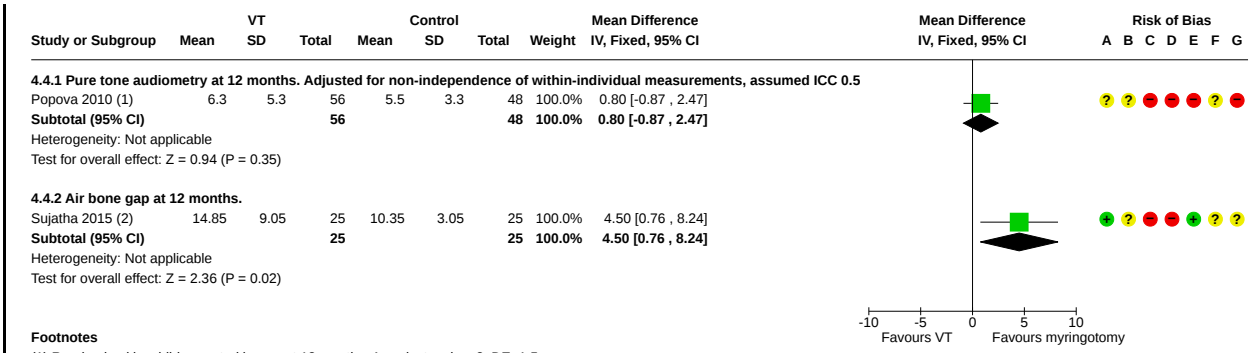
(E) Incomplete outcome data (attrition bias)

(F) Selective reporting (reporting bias)

(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 3: Mean final hearing threshold, randomised by ear (short-term)

Analysis 4.4



Footnotes

(1) Randomised by child, reported by ear at 12 months. Ave cluster size=2. DE=1.5

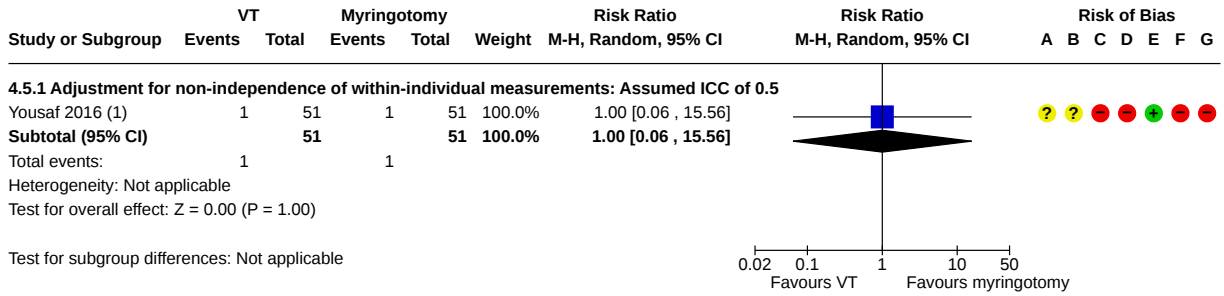
(2) Randomised by child. Both ears assessed. Correlation coefficient assumed to be 0.5.

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 4: Mean final hearing threshold (medium-term)

Analysis 4.5



Footnotes

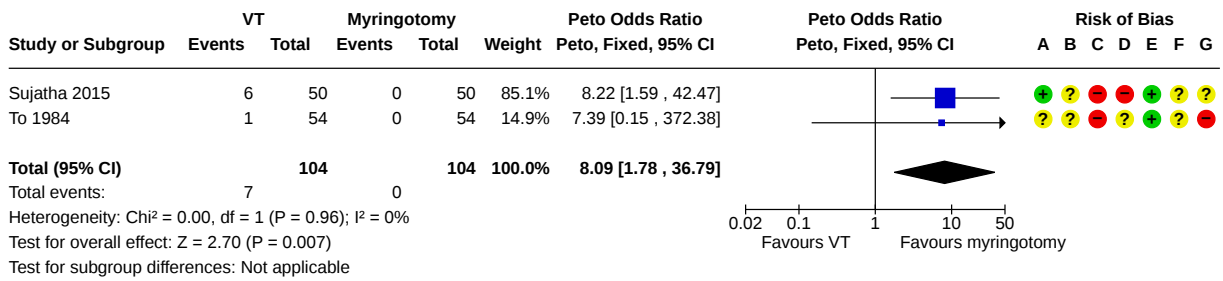
(1) 6 months. VT versus laser myringotomy. Randomised by child, reported by ears. Ave cluster size=1.66. DE=1.33

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 5: Adverse event: persistent perforation (medium-term)

Analysis 4.6



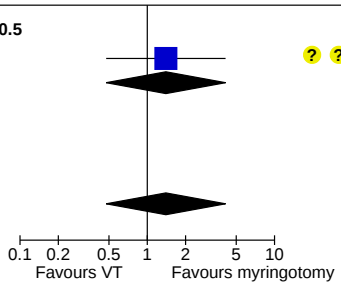
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 6: Adverse event: persistent perforation cold-steel myringotomy (medium-term)

Analysis 4.7

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
4.7.1 Adjusted for non-independence of within-individual measurements. Assumed ICC of 0.5							
Yousaf 2016 (1)	7	51	5	51	100.0%	1.40 [0.48 , 4.12]	? ? − − + − −
Subtotal (95% CI)		**51**		**51**	**100.0%**	**1.40 [0.48 , 4.12]**	
Total events:	7		5				
Heterogeneity: Not applicable							
Test for overall effect: Z = 0.61 (P = 0.54)							
Total (95% CI)		**51**		**51**	**100.0%**	**1.40 [0.48 , 4.12]**	
Total events:	7		5				
Heterogeneity: Not applicable							
Test for overall effect: Z = 0.61 (P = 0.54)							
Test for subgroup differences: Not applicable							



Footnotes

(1) VT versus laser myringotomy at 30 days. Reported by ear. Ave cluster size=1.66. DE=1.33

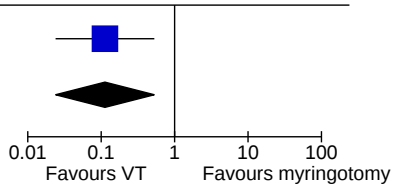
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 7: Persistence of OME: VT versus laser myringotomy (short-term)

Analysis 4.8

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Weight	Peto Odds Ratio Peto, Fixed, 95% CI
Ruckley 1988 (1)	0	36	7	36	100.0%	0.11 [0.02 , 0.53]
Total (95% CI)		**36**		**36**	**100.0%**	**0.11 [0.02 , 0.53]**
Total events:	0		7			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.77 (P = 0.006)						
Test for subgroup differences: Not applicable						



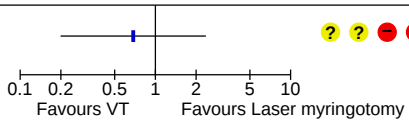
Footnotes

(1) VT versus thermal myringotomy at 3 months. No adjustment for within-individual correlation as zero events in one arm.

Comparison 4: VT versus myringotomy , Outcome 8: Persistence of OME: VT versus thermal myringotomy, randomised by ear (short-term)

Analysis 4.9

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
Popova 2010 (1)	4	42	5	36	0.69 [0.20 , 2.36]	? ? − − − ? −



Footnotes

(1) VT + Ad versus Myringotomy + Ad at 12 months.

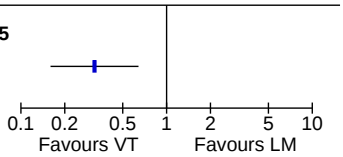
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 9: Persistence of OME: VT versus cold-steel myringotomy (medium-term)

Analysis 4.10

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Risk Ratio M-H, Random, 95% CI
4.10.1 Adjusted for non-independence of within-participant measurements: Assumed ICC of 0.5					
Yousaf 2016 (1)	8	51	25	51	0.32 [0.16 , 0.64]



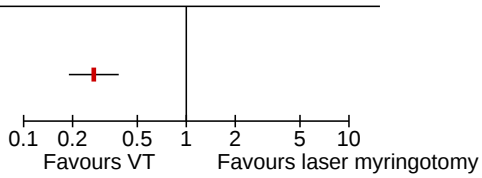
Footnotes

(1) At 6 months. Reported by ear. Ave cluster size=1.66. DE=1.33

Comparison 4: VT versus myringotomy , Outcome 10: Persistence of OME: VT versus laser myringotomy (medium-term)

Analysis 4.11

Study or Subgroup	log[Odds Ratio]	SE	Odds Ratio IV, Random, 95% CI
4.11.1 Correlation coefficient of 0.5 assumed			
Koopman 2004 (1)	-1.309333	0.18	0.27 [0.19 , 0.38]



Footnotes

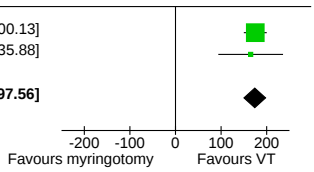
(1) Randomised by ear. Non-paired data. At 6 months.

Comparison 4: VT versus myringotomy , Outcome 11: Persistence of OME: VT versus laser myringotomy, randomised by ear (medium-term)

Analysis 4.12

Study or Subgroup	VT Mean [Days]	VT SD [Days]	VT Total	Myringotomy Mean [Days]	Myringotomy SD [Days]	Myringotomy Total	Weight	Mean Difference IV, Random, 95% CI [Days]
Gates 1989 (1)	230.6	128.701702	105	55.6	22.404741	76	88.9%	175.00 [149.87 , 200.13]
Gates 1989 (2)	262.9	164.989052	112	98	320.295279	96	11.1%	164.90 [93.92 , 235.88]
Total (95% CI)			**217**			**172**	**100.0%**	**173.88 [150.19 , 197.56]**

Heterogeneity: Tau² = 0.00; Chi² = 0.07, df = 1 (P = 0.79); I² = 0%
Test for overall effect: Z = 14.39 (P < 0.00001)
Test for subgroup differences: Not applicable



Footnotes

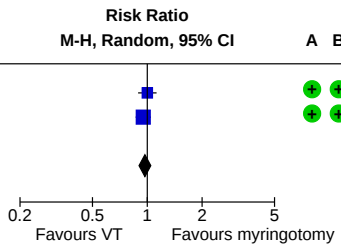
(1) VT versus M
(2) VT + Ad versus M + Ad.

Comparison 4: VT versus myringotomy , Outcome 12: Persistence of OME: mean days to first recurrence

Analysis 4.13

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Weight	Risk Ratio M-H, Random, 95% CI
Gates 1989 (1)	102	125	106	130	40.5%	1.00 [0.89 , 1.12]
Gates 1989 (2)	110	129	96	107	59.5%	0.95 [0.86 , 1.05]
Total (95% CI)		**254**		**237**	**100.0%**	**0.97 [0.90 , 1.05]**
Total events:	212		202			

Heterogeneity: Tau² = 0.00; Chi² = 0.47, df = 1 (P = 0.50); I² = 0%
Test for overall effect: Z = 0.79 (P = 0.43)
Test for subgroup differences: Not applicable



Footnotes

(1) VT + Ad versus M + Ad, 2 years follow-up.
(2) VT versus M, 2 years follow-up.

Risk of bias legend

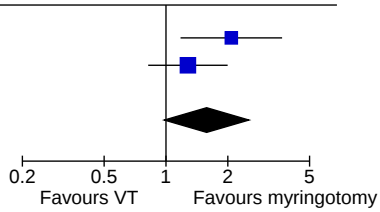
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 13: Persistence of OME (long-term)

Analysis 4.14

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Weight	Risk Ratio M-H, Random, 95% CI
Gates 1989 (1)	30	125	15	130	43.1%	2.08 [1.18 , 3.67]
Gates 1989 (2)	37	129	24	107	56.9%	1.28 [0.82 , 2.00]
Total (95% CI)		**254**		**237**	**100.0%**	**1.58 [0.98 , 2.53]**
Total events:	67		39			

Heterogeneity: Tau² = 0.05; Chi² = 1.75, df = 1 (P = 0.19); I² = 43%
Test for overall effect: Z = 1.89 (P = 0.06)
Test for subgroup differences: Not applicable

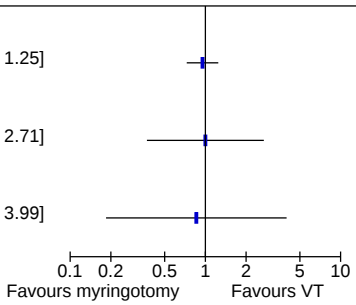


Footnotes
(1) VT + Ad versus M + Ad. Purulent ororrhoea with or without VT in place, over 2 years.
(2) VT versus M. Purulent ororrhoea with or without VT in place, over 2 years.

Comparison 4: VT versus myringotomy , Outcome 14: Adverse events: otorrhoea (long-term)

Analysis 4.15

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
4.15.1 Zero episodes						
Popova 2010 (1)	30	42	27	36	0.95 [0.73 , 1.25]	? ? − − − ? −
4.15.2 One episode						
Popova 2010 (1)	7	42	6	36	1.00 [0.37 , 2.71]	? ? − − − ? −
4.15.3 Two episodes						
Popova 2010 (1)	3	42	3	36	0.86 [0.18 , 3.99]	? ? − − − ? −



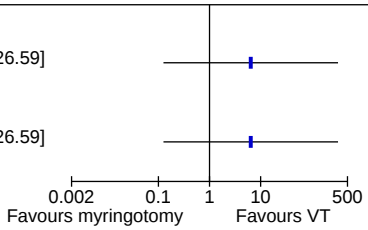
Footnotes
(1) VT + Ad versus M + Ad.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 15: Zero, one or two episodes of AOM in 12 months

Analysis 4.16

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Peto Odds Ratio Peto, Fixed, 95% CI	Risk of Bias A B C D E F G
4.16.1 Three episodes						
Popova 2010 (1)	1	42	0	36	6.41 [0.13 , 326.59]	? ? − − − ? −
4.16.2 Four or more episodes						
Popova 2010	1	42	0	36	6.41 [0.13 , 326.59]	? ? − − − ? −



Footnotes
(1) VT + Ad versus M + Ad.

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 16: Three or more episodes of AOM in 12 months

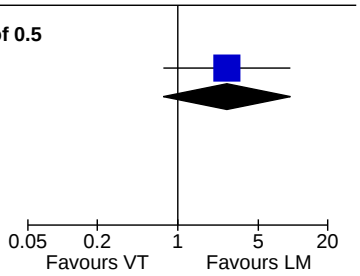
Analysis 4.17

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Weight	Risk Ratio M-H, Random, 95% CI
4.17.1 Adjusted for non-independence of within-individual measurements. Assumed ICC of 0.5						
Yousaf 2016 (1)	8	51	3	51	100.0%	2.67 [0.75 , 9.48]
Subtotal (95% CI)		**51**		**51**	**100.0%**	**2.67 [0.75 , 9.48]**
Total events:	8		3			

Heterogeneity: Not applicable

Test for overall effect: Z = 1.52 (P = 0.13)

Test for subgroup differences: Not applicable



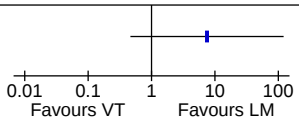
Footnotes

(1) At 6 months. Reported by ear. Ave cluster size=1.66. DE=1.33

Comparison 4: VT versus myringotomy , Outcome 17: Adverse event: retraction of TM: VT versus laser myringotomy (medium-term)

Analysis 4.18

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Peto Odds Ratio Peto, Fixed, 95% CI	Risk of Bias A B C D E F G
Yousaf 2016 (1)	2	68	0	68	7.50 [0.46 , 121.15]	? ? − − + − −



Footnotes

(1) At 6 months. Reported by ear. No adjustment for potential clustering effect as zero events for LM.

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 18: Adverse event: hypertrophic scar of TM: VT versus laser myringotomy (medium-term)

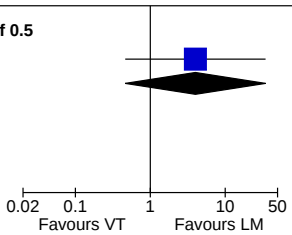
Analysis 4.19

Study or Subgroup	VT Events	VT Total	Myringotomy Events	Myringotomy Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk of Bias A B C D E F G
4.19.1 Adjusted for non-independence of within-individual measurements: assumed ICC of 0.5							
Yousaf 2016 (1)	4	51	1	51	100.0%	4.00 [0.46 , 34.57]	? ? − − + − −
Subtotal (95% CI)		**51**		**51**	**100.0%**	**4.00 [0.46 , 34.57]**	
Total events:	4		1				

Heterogeneity: Not applicable

Test for overall effect: Z = 1.26 (P = 0.21)

Test for subgroup differences: Not applicable



Footnotes

(1) At 6 months. Reported by ear. Ave cluster size=1.66. DE=1.33

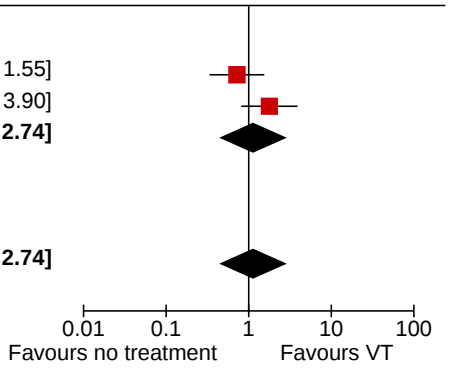
Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 4: VT versus myringotomy , Outcome 19: Adverse event: otorrhoea: VT versus laser myringotomy (medium-term)

Analysis 5.1

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
5.1.1 Sensitivity analysis: normal defined as <15dB. CC=0.3				
Dempster 1993 (1)	-0.328504	0.39	50.5%	0.72 [0.34 , 1.55]
Dempster 1993 (2)	0.576613	0.4	49.5%	1.78 [0.81 , 3.90]
Subtotal (95% CI)			**100.0%**	**1.13 [0.46 , 2.74]**
Heterogeneity: Tau² = 0.25; Chi² = 2.62, df = 1 (P = 0.11); I² = 62%				
Test for overall effect: Z = 0.26 (P = 0.79)				
Total (95% CI)			**100.0%**	**1.13 [0.46 , 2.74]**
Heterogeneity: Tau² = 0.25; Chi² = 2.62, df = 1 (P = 0.11); I² = 62%				
Test for overall effect: Z = 0.26 (P = 0.79)				
Test for subgroup differences: Not applicable				



Footnotes

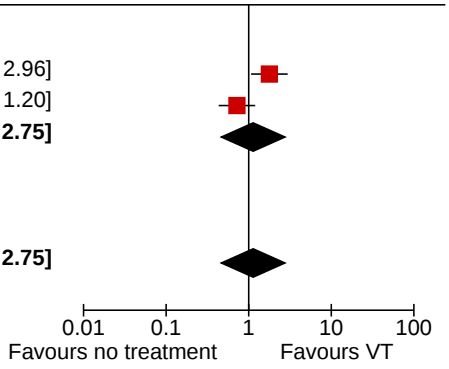
(1) Adenoidectomy and unilateral VT versus adenoidectomy only at 12 months

(2) Unilateral VT versus no treatment at 12 months

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 1: Sensitivity analysis: Return to normal hearing, randomised by ear (medium-term). CC 0.3

Analysis 5.2

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
5.2.1 Sensitivity analysis: normal defined as <15dB. CC=0.7				
Dempster 1993 (1)	0.576613	0.26	50.0%	1.78 [1.07 , 2.96]
Dempster 1993 (2)	-0.328504	0.26	50.0%	0.72 [0.43 , 1.20]
Subtotal (95% CI)			**100.0%**	**1.13 [0.47 , 2.75]**
Heterogeneity: Tau² = 0.34; Chi² = 6.06, df = 1 (P = 0.01); I² = 83%				
Test for overall effect: Z = 0.27 (P = 0.78)				
Total (95% CI)			**100.0%**	**1.13 [0.47 , 2.75]**
Heterogeneity: Tau² = 0.34; Chi² = 6.06, df = 1 (P = 0.01); I² = 83%				
Test for overall effect: Z = 0.27 (P = 0.78)				
Test for subgroup differences: Not applicable				



Footnotes

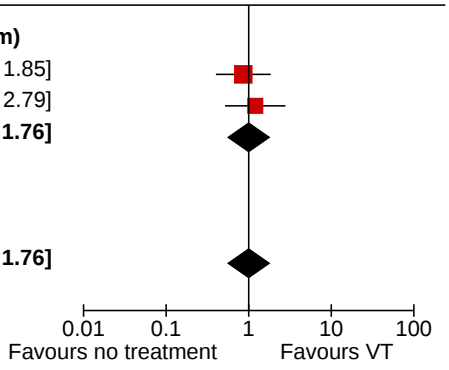
(1) Unilateral VT versus no treatment at 12 months

(2) Adenoidectomy and unilateral VT versus adenoidectomy only at 12 months

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 2: Sensitivity analysis. Return to normal hearing, randomised by ear (medium-term). CC 0.7

Analysis 5.3

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
5.3.1 Sensitivity analysis: normal defined as <25dB. CC=0.5 (medium-term)				
Dempster 1993 (1)	-0.150823	0.39	54.9%	0.86 [0.40 , 1.85]
Dempster 1993 (2)	0.182322	0.43	45.1%	1.20 [0.52 , 2.79]
Subtotal (95% CI)			**100.0%**	**1.00 [0.57 , 1.76]**
Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 1 (P = 0.57); I² = 0%				
Test for overall effect: Z = 0.00 (P = 1.00)				
Total (95% CI)			**100.0%**	**1.00 [0.57 , 1.76]**
Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 1 (P = 0.57); I² = 0%				
Test for overall effect: Z = 0.00 (P = 1.00)				
Test for subgroup differences: Not applicable				



Footnotes

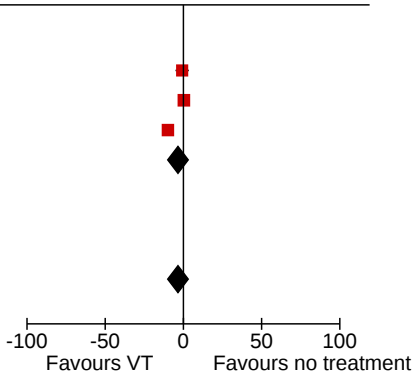
(1) Unilateral VT versus no treatment at 12 months

(2) Adenoidectomy and unilateral VT versus adenoidectomy only at 12 months

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 3: Sensitivity analysis. Return to normal hearing, randomised by ear (medium-term). Normal defined as <25dB. CC=0.5

Analysis 5.4

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI
5.4.1 Sensitivity analysis: correlation coefficient = 0.3				
Dempster 1993 (1)	-0.8	2.17	32.2%	-0.80 [-5.05 , 3.45]
Dempster 1993 (2)	0.3	1.63	34.2%	0.30 [-2.89 , 3.49]
Maw 1983 (1)	-9.9	1.83	33.5%	-9.90 [-13.49 , -6.31]
Subtotal (95% CI)			**100.0%**	**-3.47 [-10.01 , 3.06]**
Heterogeneity: Tau² = 29.85; Chi² = 19.20, df = 2 (P < 0.0001); I² = 90%				
Test for overall effect: Z = 1.04 (P = 0.30)				
Total (95% CI)			**100.0%**	**-3.47 [-10.01 , 3.06]**
Heterogeneity: Tau² = 29.85; Chi² = 19.20, df = 2 (P < 0.0001); I² = 90%				
Test for overall effect: Z = 1.04 (P = 0.30)				
Test for subgroup differences: Not applicable				



Footnotes

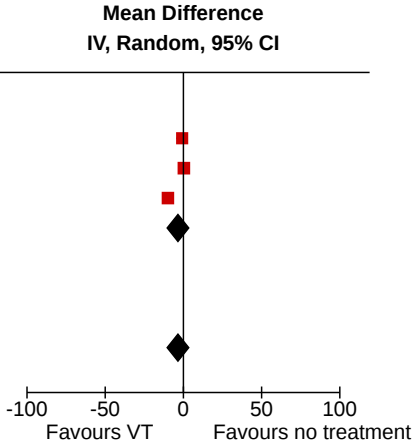
(1) Unilateral VT versus nil at 12 months

(2) Ad + unilateral VT versus ad only at 12 months

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 4: Sensitivity analysis. Mean final hearing threshold, randomised by ear (medium-term). CC0.3

Analysis 5.5

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI
5.5.1 Sensitivity analysis: correlation coefficient = 0.7				
Dempster 1993 (1)	-0.8	1.84	32.1%	-0.80 [-4.41 , 2.81]
Dempster 1993 (2)	0.3	1.07	34.1%	0.30 [-1.80 , 2.40]
Maw 1983 (1)	-9.9	1.26	33.7%	-9.90 [-12.37 , -7.43]
Subtotal (95% CI)			**100.0%**	**-3.49 [-10.37 , 3.38]**
Heterogeneity: Tau² = 34.88; Chi² = 40.56, df = 2 (P < 0.00001); I² = 95%				
Test for overall effect: Z = 1.00 (P = 0.32)				
Total (95% CI)			**100.0%**	**-3.49 [-10.37 , 3.38]**
Heterogeneity: Tau² = 34.88; Chi² = 40.56, df = 2 (P < 0.00001); I² = 95%				
Test for overall effect: Z = 1.00 (P = 0.32)				
Test for subgroup differences: Not applicable				



Footnotes

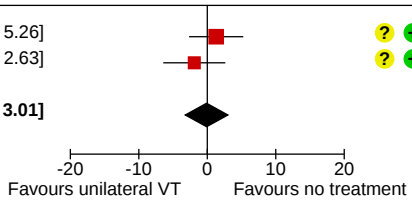
(1) Unilateral VT versus nil at 12 months

(2) Ad + unilateral VT versus ad only at 12 months

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 5: Sensitivity analysis. Mean final hearing threshold, randomised by ear (medium-term). CC0.7

Analysis 5.6

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Dempster 1993 (1)	1.3	2.02	56.1%	1.30 [-2.66 , 5.26]		? + − ? ? ? ?
Dempster 1993 (2)	-1.9	2.31	43.9%	-1.90 [-6.43 , 2.63]		? + − ? ? ? ?
Total (95% CI)			**100.0%**	**-0.10 [-3.22 , 3.01]**		



Heterogeneity: Tau² = 0.41; Chi² = 1.09, df = 1 (P = 0.30); I² = 8%
Test for overall effect: Z = 0.07 (P = 0.95)
Test for subgroup differences: Not applicable

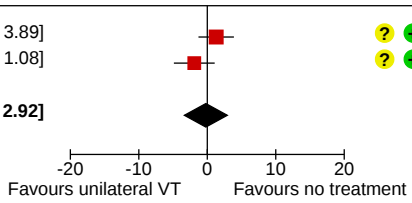
Footnotes
(1) Adenoidectomy plus unilateral VT versus adenoidectomy only. CC=0.3
(2) Unilateral VT versus nil. CC=0.3

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 6: Sensitivity analysis. Change in hearing threshold from baseline, randomised by ear (medium-term). CC0.3

Analysis 5.7

Study or Subgroup	Mean Difference	SE	Weight	Mean Difference IV, Random, 95% CI	Mean Difference IV, Random, 95% CI	Risk of Bias A B C D E F G
Dempster 1993 (1)	1.3	1.32	52.8%	1.30 [-1.29 , 3.89]		? + − ? ? ? ?
Dempster 1993 (2)	-1.9	1.52	47.2%	-1.90 [-4.88 , 1.08]		? + − ? ? ? ?
Total (95% CI)			**100.0%**	**-0.21 [-3.34 , 2.92]**		



Heterogeneity: Tau² = 3.09; Chi² = 2.53, df = 1 (P = 0.11); I² = 60%
Test for overall effect: Z = 0.13 (P = 0.89)
Test for subgroup differences: Not applicable

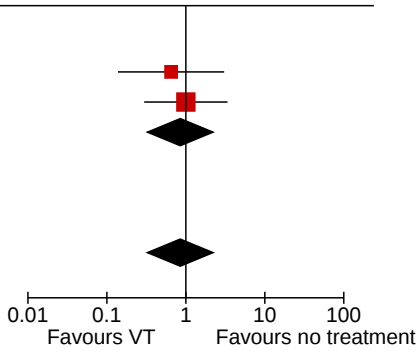
Footnotes
(1) Adenoidectomy plus unilateral VT versus adenoidectomy only. CC=0.7
(2) Unilateral VT versus nil. CC=0.7

Risk of bias legend
(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 7: Sensitivity analysis. Change in hearing threshold from baseline, randomised by ear (medium-term). CC0.7

Analysis 5.8

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
5.8.1 Sensitivity analysis: correlation coefficient 0.3				
Dempster 1993 (1)	-0.430783	0.79	38.1%	0.65 [0.14 , 3.06]
Dempster 1993 (2)	0	0.62	61.9%	1.00 [0.30 , 3.37]
Subtotal (95% CI)			**100.0%**	**0.85 [0.33 , 2.21]**
Heterogeneity: Tau² = 0.00; Chi² = 0.18, df = 1 (P = 0.67); I² = 0%				
Test for overall effect: Z = 0.34 (P = 0.74)				
Total (95% CI)			**100.0%**	**0.85 [0.33 , 2.21]**
Heterogeneity: Tau² = 0.00; Chi² = 0.18, df = 1 (P = 0.67); I² = 0%				
Test for overall effect: Z = 0.34 (P = 0.74)				
Test for subgroup differences: Not applicable				



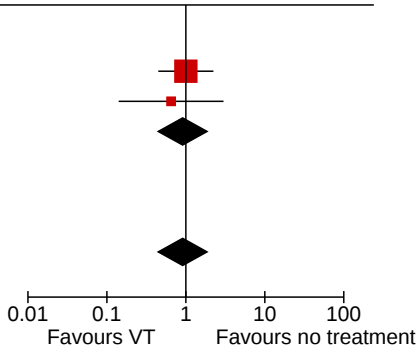
Footnotes

(1) Unilateral VT versus no treatment at 12 months.
(2) Unilateral VT + Ad versus Ad only, at 12 months.

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 8: Sensitivity analysis. Adverse event: perforation/retraction, randomised by ear (medium-term). CC=0.3

Analysis 5.9

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
5.9.1 Sensitivity analysis: correlation coefficient 0.7				
Dempster 1993 (1)	0	0.41	78.4%	1.00 [0.45 , 2.23]
Dempster 1993 (2)	-0.430783	0.78	21.6%	0.65 [0.14 , 3.00]
Subtotal (95% CI)			**100.0%**	**0.91 [0.45 , 1.86]**
Heterogeneity: Tau² = 0.00; Chi² = 0.24, df = 1 (P = 0.62); I² = 0%				
Test for overall effect: Z = 0.26 (P = 0.80)				
Total (95% CI)			**100.0%**	**0.91 [0.45 , 1.86]**
Heterogeneity: Tau² = 0.00; Chi² = 0.24, df = 1 (P = 0.62); I² = 0%				
Test for overall effect: Z = 0.26 (P = 0.80)				
Test for subgroup differences: Not applicable				



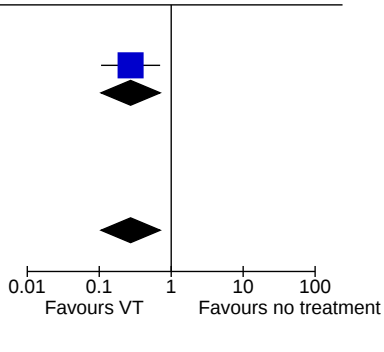
Footnotes

(1) Unilateral VT + Ad versus Ad only, at 12 months.
(2) Unilateral VT versus no treatment at 12 months.

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 9: Sensitivity analysis. Adverse event: perforation/retraction, randomised by ear (medium-term). CC=0.7

Analysis 5.10

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
5.10.1 Sensitivity analysis: assuming ICC of 1.0 (complete correlation between ears)						
Rach 1991 (1)	4	22	12	18	100.0%	0.27 [0.11 , 0.70]
Subtotal (95% CI)		**22**		**18**	**100.0%**	**0.27 [0.11 , 0.70]**
Total events:	4		12			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.70 (P = 0.007)						
Total (95% CI)		**22**		**18**	**100.0%**	**0.27 [0.11 , 0.70]**
Total events:	4		12			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.70 (P = 0.007)						
Test for subgroup differences: Not applicable						



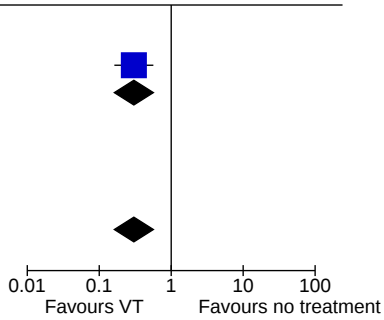
Footnotes

(1) Bilateral VT versus nil at 6 months. Analysed by ear. Average cluster size=2. DE=2.0

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 10: Sensitivity analysis. Persistence of OME: randomised by child (medium-term). ICC 1.0

Analysis 5.11

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
5.11.1 Sensitivity analysis: assuming ICC of 0.0 (no correlation between ears)						
Rach 1991 (1)	9	44	25	37	100.0%	0.30 [0.16 , 0.56]
Subtotal (95% CI)		**44**		**37**	**100.0%**	**0.30 [0.16 , 0.56]**
Total events:	9		25			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.75 (P = 0.0002)						
Total (95% CI)		**44**		**37**	**100.0%**	**0.30 [0.16 , 0.56]**
Total events:	9		25			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.75 (P = 0.0002)						
Test for subgroup differences: Not applicable						



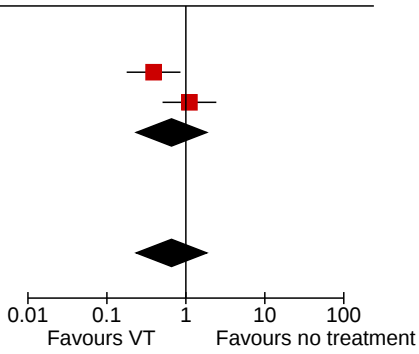
Footnotes

(1) Bilateral VT versus nil at 6 months. Analysed by ear. Average cluster size=2. DE=1

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 11: Sensitivity analysis. Persistence of OME: randomised by child (medium-term). ICC zero

Analysis 5.12

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
5.12.1 Sensitivity analysis: correlation coefficient = 0.3				
Dempster 1993 (1)	-0.941609	0.4	50.0%	0.39 [0.18 , 0.85]
Dempster 1993 (2)	0.10436	0.4	50.0%	1.11 [0.51 , 2.43]
Subtotal (95% CI)			**100.0%**	**0.66 [0.24 , 1.83]**
Heterogeneity: Tau² = 0.39; Chi² = 3.42, df = 1 (P = 0.06); I² = 71%				
Test for overall effect: Z = 0.80 (P = 0.42)				
Total (95% CI)			**100.0%**	**0.66 [0.24 , 1.83]**
Heterogeneity: Tau² = 0.39; Chi² = 3.42, df = 1 (P = 0.06); I² = 71%				
Test for overall effect: Z = 0.80 (P = 0.42)				
Test for subgroup differences: Not applicable				



Footnotes

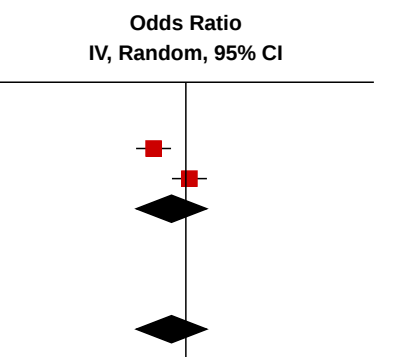
(1) Unilateral VT versus nil. Tympanometry at 12 months.

(2) Adenoidectomy plus unilateral VT versus adenoidectomy alone

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 12: Sensitivity analysis. Persistence of OME: randomised by ear (medium-term). CC 0.3

Analysis 5.13

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
5.13.1 Sensitivity analysis: correlation coefficient = 0.7				
Dempster 1993 (1)	-0.941609	0.26	50.0%	0.39 [0.23 , 0.65]
Dempster 1993 (2)	0.10436	0.26	50.0%	1.11 [0.67 , 1.85]
Subtotal (95% CI)			**100.0%**	**0.66 [0.24 , 1.83]**
Heterogeneity: Tau² = 0.48; Chi² = 8.09, df = 1 (P = 0.004); I² = 88%				
Test for overall effect: Z = 0.80 (P = 0.42)				
Total (95% CI)			**100.0%**	**0.66 [0.24 , 1.83]**
Heterogeneity: Tau² = 0.48; Chi² = 8.09, df = 1 (P = 0.004); I² = 88%				
Test for overall effect: Z = 0.80 (P = 0.42)				
Test for subgroup differences: Not applicable				



Footnotes

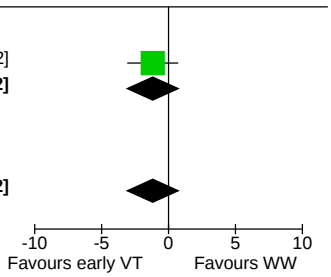
(1) Unilateral VT versus nil. Tympanometry at 12 months.

(2) Adenoidectomy plus unilateral VT versus adenoidectomy alone

Comparison 5: Sensitivity analyses: VT versus no treatment, Outcome 13: Sensitivity analysis. Persistence of OME: randomised by ear (medium-term). CC 0.7

Analysis 6.1

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.1.1 Sensitivity analysis: assuming ICC of 1.0 (complete correlation between ears)								
Velepic 2011 (1)	6.02	3.81	32	7.2	5.19	55	100.0%	-1.18 [-3.08 , 0.72]
Subtotal (95% CI)			**32**			**55**	**100.0%**	**-1.18 [-3.08 , 0.72]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.21 (P = 0.22)								
Total (95% CI)			**32**			**55**	**100.0%**	**-1.18 [-3.08 , 0.72]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.21 (P = 0.22)								
Test for subgroup differences: Not applicable								



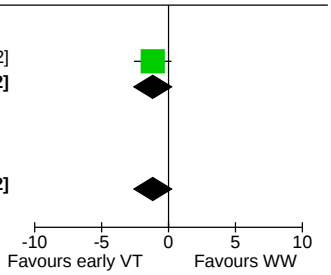
Footnotes

(1) Average cluster size=1.85. DE=1.85

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 1: Sensitivity analysis. Mean final hearing threshold (air-bone gap), randomised by child, analysed by ear (medium-term). ICC 1.0

Analysis 6.2

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.2.1 Sensitivity analysis: assuming ICC of 0.0 (no correlation between ears)								
Velepic 2011 (1)	6.02	3.81	59	7.2	5.19	101	100.0%	-1.18 [-2.58 , 0.22]
Subtotal (95% CI)			**59**			**101**	**100.0%**	**-1.18 [-2.58 , 0.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.65 (P = 0.10)								
Total (95% CI)			**59**			**101**	**100.0%**	**-1.18 [-2.58 , 0.22]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.65 (P = 0.10)								
Test for subgroup differences: Not applicable								



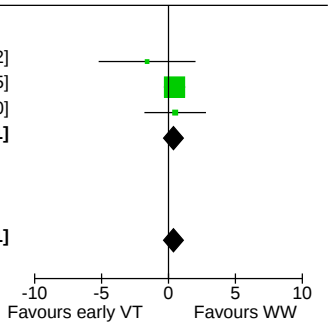
Footnotes

(1) Average cluster size=1.85. DE=1.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 2: Sensitivity analysis. Mean final hearing threshold (air-bone gap), randomised by child, analysed by ear (medium-term). ICC zero

Analysis 6.3

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.3.1 Sensitivity analysis: cc for Paradise 2007 (left and right ear data combined) of 0.3								
Maw 1999 (1)	12.7	11.5	75	14.3	10.5	67	4.1%	-1.60 [-5.22 , 2.02]
Paradise 2007 (2)	6.2	3.3	147	5.75	3.49	134	85.6%	0.45 [-0.35 , 1.25]
TARGET 2000 (3)	18.7	8.9	108	18.2	8.1	102	10.3%	0.50 [-1.80 , 2.80]
Subtotal (95% CI)			**330**			**303**	**100.0%**	**0.37 [-0.37 , 1.11]**
Heterogeneity: Tau² = 0.00; Chi² = 1.19, df = 2 (P = 0.55); I² = 0%								
Test for overall effect: Z = 0.99 (P = 0.32)								
Total (95% CI)			**330**			**303**	**100.0%**	**0.37 [-0.37 , 1.11]**
Heterogeneity: Tau² = 0.00; Chi² = 1.19, df = 2 (P = 0.55); I² = 0%								
Test for overall effect: Z = 0.99 (P = 0.32)								
Test for subgroup differences: Not applicable								



Footnotes

(1) Bilateral VT versus WW at 18 months, best ear at 4000Hz.

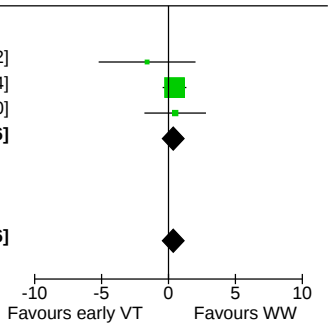
(2) At age 5. R and L ear data combined, with correction of variance. Assumed correlation coeff. of 0.3

(3) Bilateral VT versus WW at 2 years. Maximum cases available.

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 3: Sensitivity analysis. Mean final hearing threshold, randomised by child (long-term). CC for Paradise 2007 of 0.3

Analysis 6.4

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
6.4.1 Sensitivity analysis: cc for Paradise 2007 (left and right ear data combined) of 0.7								
Maw 1999 (1)	12.7	11.5	75	14.3	10.5	67	5.0%	-1.60 [-5.22 , 2.02]
Paradise 2007 (2)	6.2	3.78	147	5.75	3.8	134	82.7%	0.45 [-0.44 , 1.34]
TARGET 2000 (3)	18.7	8.9	108	18.2	8.1	102	12.3%	0.50 [-1.80 , 2.80]
Subtotal (95% CI)			**330**			**303**	**100.0%**	**0.35 [-0.45 , 1.16]**
Heterogeneity: Tau² = 0.00; Chi² = 1.18, df = 2 (P = 0.55); I² = 0%								
Test for overall effect: Z = 0.86 (P = 0.39)								
Total (95% CI)			**330**			**303**	**100.0%**	**0.35 [-0.45 , 1.16]**
Heterogeneity: Tau² = 0.00; Chi² = 1.18, df = 2 (P = 0.55); I² = 0%								
Test for overall effect: Z = 0.86 (P = 0.39)								
Test for subgroup differences: Not applicable								



Footnotes

(1) Bilateral VT versus WW at 18 months, best ear at 4000Hz.

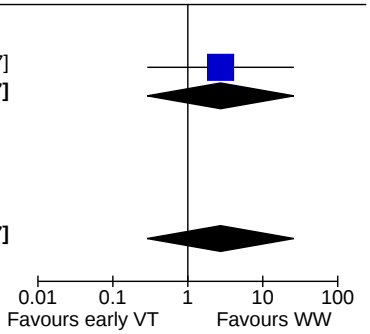
(2) At age 5. R and L ear data combined, with correction of variance. Assumed correlation coeff. of 0.7.

(3) Bilateral VT versus WW at 2 years. Maximum cases available.

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 4: Sensitivity analysis. Mean final hearing threshold, randomised by child (long-term). CC for Paradise 2007 of 0.7

Analysis 6.5

Study or Subgroup	Experimental Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
6.5.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)						
Paradise 2007 (1)	3	147	1	134	100.0%	2.73 [0.29 , 25.97]
Subtotal (95% CI)		**147**		**134**	**100.0%**	**2.73 [0.29 , 25.97]**
Total events:	3		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.88 (P = 0.38)						
Total (95% CI)		**147**		**134**	**100.0%**	**2.73 [0.29 , 25.97]**
Total events:	3		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.88 (P = 0.38)						
Test for subgroup differences: Not applicable						



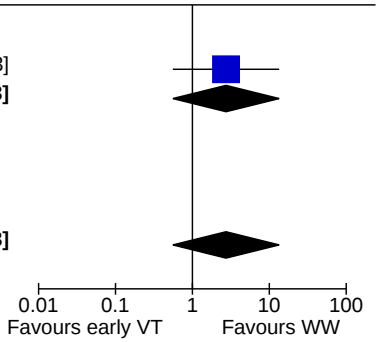
Footnotes

(1) At age 5. Analysis by ears. Each child contributed 2 data points, so average cluster size=2. DE=2.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 5: Sensitivity analysis. Persistent perforation, randomised by child (long-term). ICC 1.0

Analysis 6.6

Study or Subgroup	Experimental Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
6.6.1 Sensitivity analysis: ICC zero (no correlation between ears)						
Paradise 2007 (1)	6	294	2	268	100.0%	2.73 [0.56 , 13.43]
Subtotal (95% CI)		**294**		**268**	**100.0%**	**2.73 [0.56 , 13.43]**
Total events:	6		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.24 (P = 0.22)						
Total (95% CI)		**294**		**268**	**100.0%**	**2.73 [0.56 , 13.43]**
Total events:	6		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.24 (P = 0.22)						
Test for subgroup differences: Not applicable						



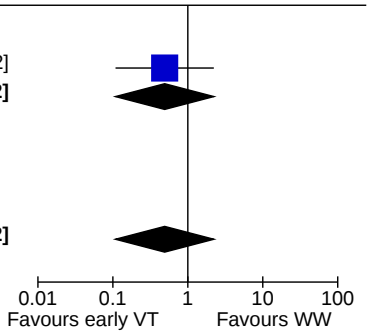
Footnotes

(1) At age 5. Analysis by ears. Each child contributed 2 data points, so average cluster size=2. DE=1.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 6: Sensitivity analysis. Persistent perforation, randomised by child (long-term). ICC zero

Analysis 6.7

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
6.7.1 Sensitivity analysis: assuming ICC of 1.0 (complete correlation between ears)						
Velepic 2011 (1)	2	32	7	55	100.0%	0.49 [0.11 , 2.22]
Subtotal (95% CI)		**32**		**55**	**100.0%**	**0.49 [0.11 , 2.22]**
Total events:	2		7			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.92 (P = 0.36)						
Total (95% CI)		**32**		**55**	**100.0%**	**0.49 [0.11 , 2.22]**
Total events:	2		7			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.92 (P = 0.36)						
Test for subgroup differences: Not applicable						



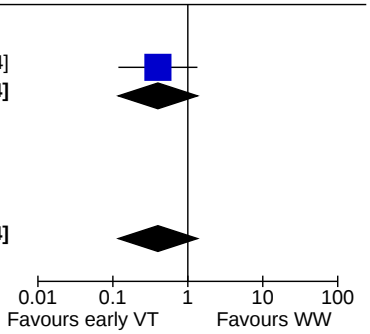
Footnotes

(1) At least 6 mo after surgery. Analysed by ear. Ave cluster size=1.85. DE=1.85

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 7: Sensitivity analysis. Persistence of OME, randomised by child, measured by otoscopy (medium-term). ICC 1.0

Analysis 6.8

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
6.8.1 Sensitivity analysis: assuming ICC of 0.0 (no correlation between ears)						
Velepic 2011 (1)	3	59	13	102	100.0%	0.40 [0.12 , 1.34]
Subtotal (95% CI)		**59**		**102**	**100.0%**	**0.40 [0.12 , 1.34]**
Total events:	3		13			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.48 (P = 0.14)						
Total (95% CI)		**59**		**102**	**100.0%**	**0.40 [0.12 , 1.34]**
Total events:	3		13			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.48 (P = 0.14)						
Test for subgroup differences: Not applicable						



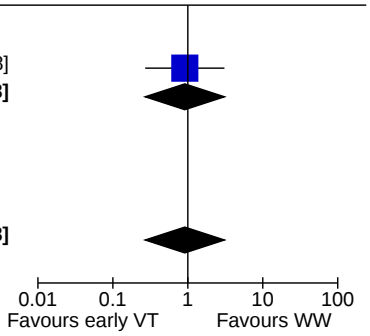
Footnotes

(1) At least 6 mo after surgery. Analysed by ear. Ave cluster size=1.85. DE=1.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 8: Sensitivity analysis. Persistence of OME, randomised by child, measured by otoscopy (medium-term). ICC=zero

Analysis 6.9

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
6.9.1 Sensitivity analysis: ICC 1.0 (full correlation between ears)						
Paradise 2007 (1)	5	147	5	134	100.0%	0.91 [0.27 , 3.08]
Subtotal (95% CI)		**147**		**134**	**100.0%**	**0.91 [0.27 , 3.08]**
Total events:	5		5			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.15 (P = 0.88)						
Total (95% CI)		**147**		**134**	**100.0%**	**0.91 [0.27 , 3.08]**
Total events:	5		5			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.15 (P = 0.88)						
Test for subgroup differences: Not applicable						



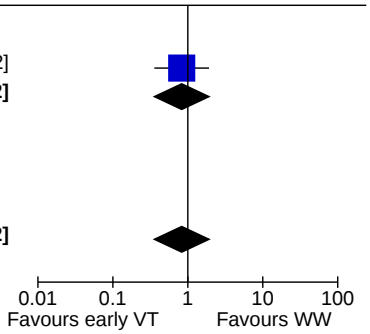
Footnotes

(1) At age 5. Assessed using otomicrosopy. Ave cluster size=2. DE=2.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 9: Sensitivity analysis. Tympanosclerosis (long term). ICC=1.0

Analysis 6.10

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
6.10.1 Sensitivity analysis ICC zero (no correlation between ears)						
Paradise 2007 (1)	10	294	11	268	100.0%	0.83 [0.36 , 1.92]
Subtotal (95% CI)		**294**		**268**	**100.0%**	**0.83 [0.36 , 1.92]**
Total events:	10		11			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.44 (P = 0.66)						
Total (95% CI)		**294**		**268**	**100.0%**	**0.83 [0.36 , 1.92]**
Total events:	10		11			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.44 (P = 0.66)						
Test for subgroup differences: Not applicable						



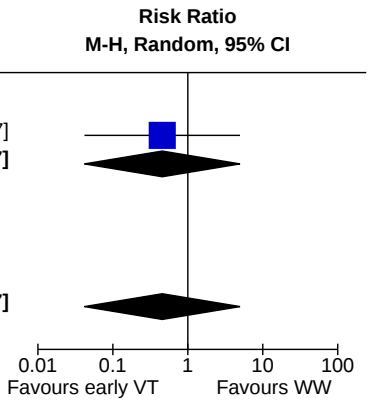
Footnotes

(1) At age 5. Assessed using otomicrosopy. Ave cxluster size=2. DE=1

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 10: Sensitivity analysis. Tympanosclerosis (long term). ICC=zero

Analysis 6.11

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
6.11.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)						
Paradise 2007 (1)	1	147	2	134	100.0%	0.46 [0.04 , 4.97]
Subtotal (95% CI)		**147**		**134**	**100.0%**	**0.46 [0.04 , 4.97]**
Total events:	1		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.64 (P = 0.52)						
Total (95% CI)		**147**		**134**	**100.0%**	**0.46 [0.04 , 4.97]**
Total events:	1		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.64 (P = 0.52)						
Test for subgroup differences: Not applicable						



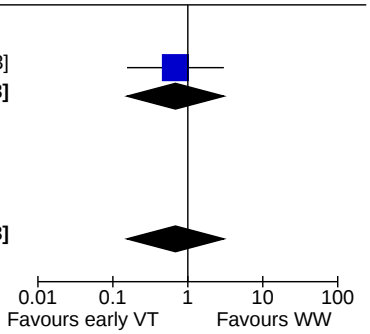
Footnotes

(1) assessed using otomicrosopy. Ave cluster size=2. DE=2.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 11: Sensitivity analysis. Adverse event: fibrosis (long term). ICC=1.0

Analysis 6.12

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
6.12.1 Sensitivity analysis: ICC zero (no correlation between ears)						
Paradise 2007 (1)	3	294	4	268	100.0%	0.68 [0.15 , 3.03]
Subtotal (95% CI)		**294**		**268**	**100.0%**	**0.68 [0.15 , 3.03]**
Total events:	3		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.50 (P = 0.62)						
Total (95% CI)		**294**		**268**	**100.0%**	**0.68 [0.15 , 3.03]**
Total events:	3		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.50 (P = 0.62)						
Test for subgroup differences: Not applicable						



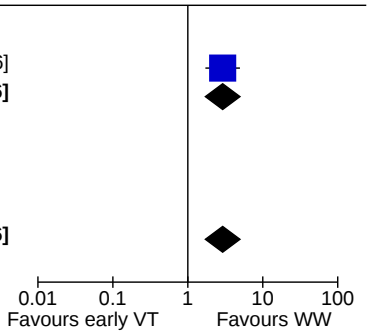
Footnotes

(1) assessed using otomicrosopy. Ave cluster size=2. DE=1.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 12: Sensitivity analysis. Adverse event: fibrosis (long term). ICC=zero

Analysis 6.13

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
6.13.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)						
Paradise 2007 (1)	48	147	15	134	100.0%	2.92 [1.72 , 4.96]
Subtotal (95% CI)		**147**		**134**	**100.0%**	**2.92 [1.72 , 4.96]**
Total events:	48		15			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.96 (P < 0.0001)						
Total (95% CI)		**147**		**134**	**100.0%**	**2.92 [1.72 , 4.96]**
Total events:	48		15			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.96 (P < 0.0001)						
Test for subgroup differences: Not applicable						



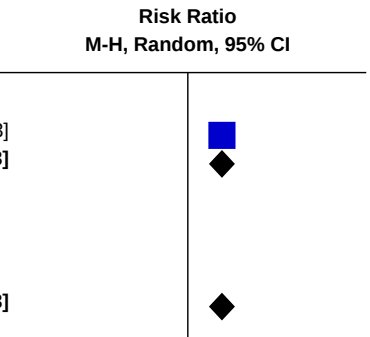
Footnotes

(1) Age 5 years. assessed using otomicrosopy. Ave cluster size=2. DE=2.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 13: Sensitivity analysis. Segmental atrophy (long term). ICC=1.0

Analysis 6.14

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
6.14.1 Sensitivity analysis: ICC zero (no correlation between ears)						
Paradise 2007 (1)	97	294	31	268	100.0%	2.85 [1.97 , 4.13]
Subtotal (95% CI)		**294**		**268**	**100.0%**	**2.85 [1.97 , 4.13]**
Total events:	97		31			
Heterogeneity: Not applicable						
Test for overall effect: Z = 5.57 (P < 0.00001)						
Total (95% CI)		**294**		**268**	**100.0%**	**2.85 [1.97 , 4.13]**
Total events:	97		31			
Heterogeneity: Not applicable						
Test for overall effect: Z = 5.57 (P < 0.00001)						
Test for subgroup differences: Not applicable						



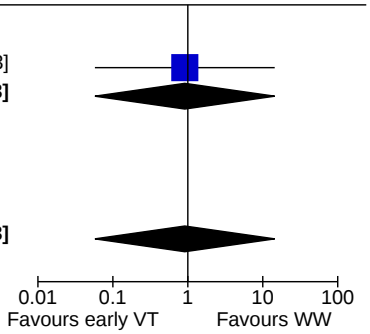
Footnotes

(1) Age 5 years. assessed using otomicrosopy. Ave cluster size=2. DE=1.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 14: Sensitivity analysis. Segmental atrophy (long term). ICC=zero

Analysis 6.15

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
6.15.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)						
Paradise 2007 (1)	1	147	1	134	100.0%	0.91 [0.06 , 14.43]
Subtotal (95% CI)		**147**		**134**	**100.0%**	**0.91 [0.06 , 14.43]**
Total events:	1		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.07 (P = 0.95)						
Total (95% CI)		**147**		**134**	**100.0%**	**0.91 [0.06 , 14.43]**
Total events:	1		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.07 (P = 0.95)						
Test for subgroup differences: Not applicable						



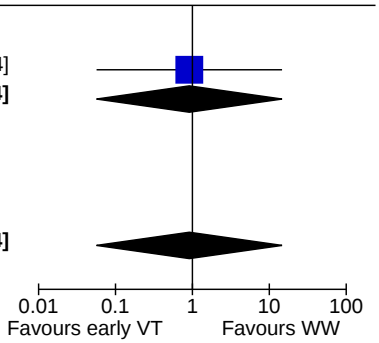
Footnotes

(1) Reported by ears. Assessed using otomicrosopy. Ave cluster size=2. DE=2.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 15: Sensitivity analysis. Retraction pocket with other abnormality (long term). ICC=1.0

Analysis 6.16

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Odds Ratio M-H, Fixed, 95% CI
6.16.1 Sensitivity analysis: ICC zero (no correlation between ears)						
Paradise 2007 (1)	1	294	1	268	100.0%	0.91 [0.06 , 14.64]
Subtotal (95% CI)		**294**		**268**	**100.0%**	**0.91 [0.06 , 14.64]**
Total events:	1		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.07 (P = 0.95)						
Total (95% CI)		**294**		**268**	**100.0%**	**0.91 [0.06 , 14.64]**
Total events:	1		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.07 (P = 0.95)						
Test for subgroup differences: Not applicable						



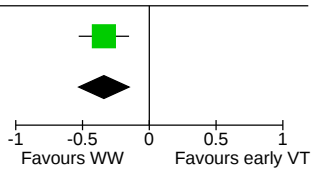
Footnotes

(1) Reported by ears. Assessed using otomicrosopy. Ave cluster size=2. DE=1.0

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 16: Sensitivity analysis. Retraction pocket with other abnormality (long term). ICC=zero

Analysis 6.17

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
Rovers 2000 (1)	5.88	0.689	84	6.22	0.54	81	100.0%	-0.34 [-0.53 , -0.15]
Total (95% CI)			**84**			**81**	**100.0%**	**-0.34 [-0.53 , -0.15]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 3.53 (P = 0.0004)								
Test for subgroup differences: Not applicable								



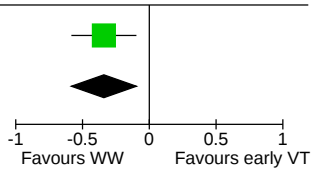
Footnotes

(1) At 12 months. Combined means across five domains, with correction of variance. Assumed correlation coeff. of 0.3. Higher = better.

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 17: Sensitivity analysis. Parent-child interaction: Erickson child scale (medium-term). CC0.3

Analysis 6.18

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
Rovers 2000 (1)	5.88	0.896	84	6.22	0.69	81	100.0%	-0.34 [-0.58 , -0.10]
Total (95% CI)			**84**			**81**	**100.0%**	**-0.34 [-0.58 , -0.10]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 2.74 (P = 0.006)								
Test for subgroup differences: Not applicable								

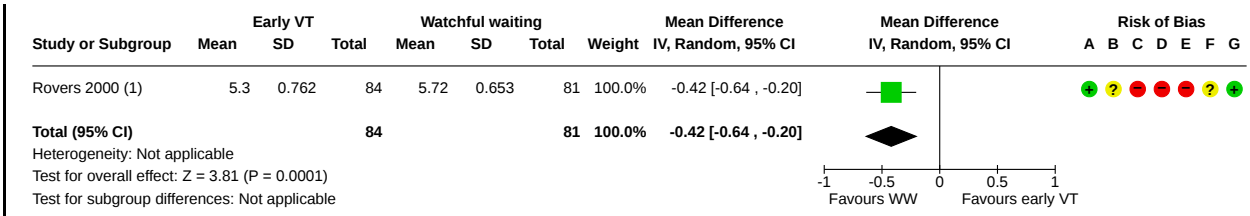


Footnotes

(1) At 12 months. Combined means across five domains, with correction of variance. Assumed correlation coeff. of 0.7. Higher = better.

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 18: Sensitivity analysis. Parent-child interaction: Erickson child scale (medium-term). CC0.7

Analysis 6.19



Footnotes

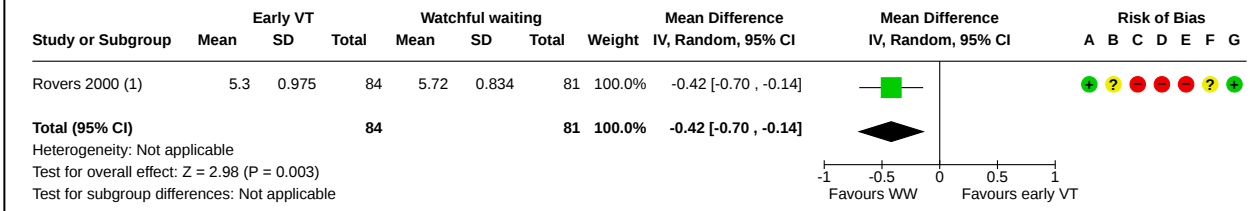
(1) At 12 months. Combined means across five domains, with correction of variance. Assumed correlation coeff. of 0.3. Higher = better.

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 19: Sensitivity analysis. Parent-child interaction: Erickson parent scale (medium-term). CC0.3

Analysis 6.20



Footnotes

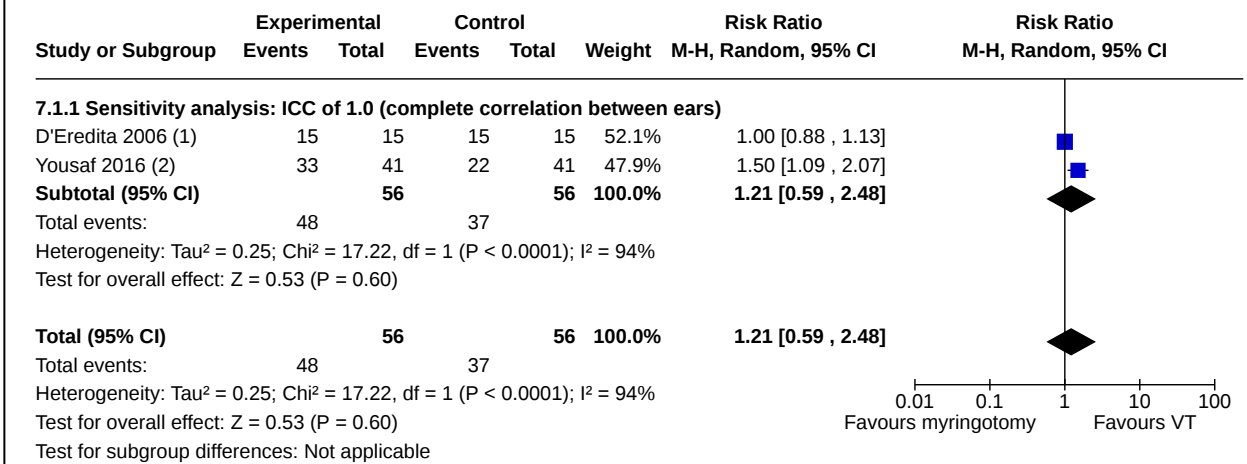
(1) At 12 months. Combined means across five domains, with correction of variance. Assumed correlation coeff. of 0.7. Higher = better.

Risk of bias legend

(A) Random sequence generation (selection bias)
(B) Allocation concealment (selection bias)
(C) Blinding of participants and personnel (performance bias)
(D) Blinding of outcome assessment (detection bias)
(E) Incomplete outcome data (attrition bias)
(F) Selective reporting (reporting bias)
(G) Other bias

Comparison 6: Sensitivity analyses: Early VT versus watchful waiting, Outcome 20: Sensitivity analysis. Parent-child interaction: Erickson parent scale (medium-term). CC=0.7

Analysis 7.1



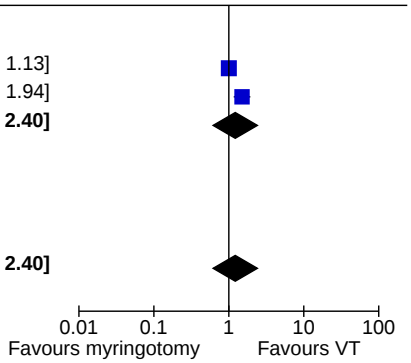
Footnotes

(1) Bilateral VT versus bilateral laser myringotomy at one year follow-up.
(2) At 6 months. Unilateral or bilateral treatment in each group. Reported by ear. Ave cluster size=1.66. DE=1.66

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 1: Sensitivity analysis. Hearing returned to normal: VT versus laser myringotomy (medium-term). ICC=1.0

Analysis 7.2

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
7.2.1 Sensitivity analysis: ICC of zero (no correlation between ears)						
D'Eredita 2006 (1)	15	15	15	15	51.3%	1.00 [0.88 , 1.13]
Yousaf 2016 (2)	54	68	36	68	48.7%	1.50 [1.16 , 1.94]
Subtotal (95% CI)		**83**		**83**	**100.0%**	**1.22 [0.62 , 2.40]**
Total events:	69		51			
Heterogeneity: Tau² = 0.23; Chi² = 22.94, df = 1 (P < 0.00001); I² = 96%						
Test for overall effect: Z = 0.57 (P = 0.57)						
Total (95% CI)		**83**		**83**	**100.0%**	**1.22 [0.62 , 2.40]**
Total events:	69		51			
Heterogeneity: Tau² = 0.23; Chi² = 22.94, df = 1 (P < 0.00001); I² = 96%						
Test for overall effect: Z = 0.57 (P = 0.57)						
Test for subgroup differences: Not applicable						



Footnotes

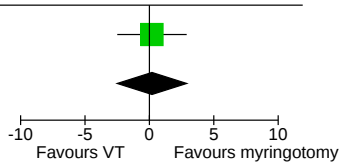
(1) Bilateral VT versus bilateral laser myringotomy at one year follow-up.

(2) At 6 months. Unilateral or bilateral treatment in each group. Reported by ear. Ave cluster size=1.66. DE=1

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 2: Sensitivity analysis. Hearing returned to normal: VT versus laser myringotomy (medium-term). ICC=zero

Analysis 7.3

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
Popova 2010 (1)	14.1	6	42	13.9	6.1	36	100.0%	0.20 [-2.50 , 2.90]
Total (95% CI)			**42**			**36**	**100.0%**	**0.20 [-2.50 , 2.90]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.15 (P = 0.88)								
Test for subgroup differences: Not applicable								



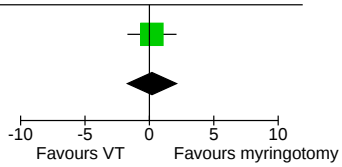
Footnotes

(1) 1 month. Ave cluster size=2. DE=2

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 3: Sensitivity analysis. Mean final hearing threshold, randomised by child (short-term). ICC 1.0

Analysis 7.4

Study or Subgroup	Experimental Mean	Experimental SD	Experimental Total	Control Mean	Control SD	Control Total	Weight	Mean Difference IV, Random, 95% CI
Popova 2010 (1)	14.1	6	84	13.9	6.1	72	100.0%	0.20 [-1.71 , 2.11]
Total (95% CI)			**84**			**72**	**100.0%**	**0.20 [-1.71 , 2.11]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.21 (P = 0.84)								
Test for subgroup differences: Not applicable								



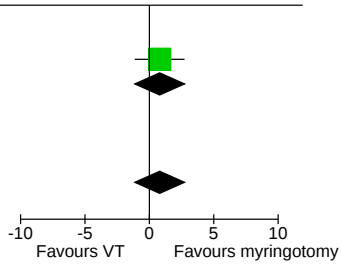
Footnotes

(1) 1 month. Ave cluster size=2. DE=1

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 4: Sensitivity analysis. Mean final hearing threshold, randomised by child (short-term). ICC=zero

Analysis 7.5

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.5.1 Sensitivity analysis: ICC 1.0 (complete correlation between ears)								
Popova 2010 (1)	6.3	5.3	42	5.5	3.3	36	100.0%	0.80 [-1.13 , 2.73]
Subtotal (95% CI)			**42**			**36**	**100.0%**	**0.80 [-1.13 , 2.73]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.81 (P = 0.42)								
Total (95% CI)			**42**			**36**	**100.0%**	**0.80 [-1.13 , 2.73]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 0.81 (P = 0.42)								
Test for subgroup differences: Not applicable								



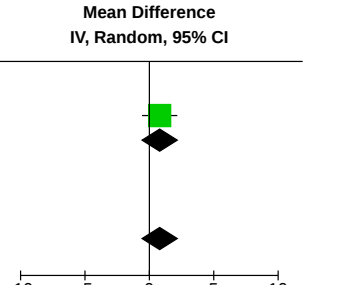
Footnotes

(1) Randomised by child, reported by ear at 12 months. Ave cluster size=2. DE=2

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 5: Sensitivity analysis. Mean final hearing threshold (medium-term). ICC=1.0

Analysis 7.6

Study or Subgroup	Experimental Mean	SD	Total	Control Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
7.6.1 Sensitivity analysis: ICC zero (no correlation between ears)								
Popova 2010 (1)	6.3	5.3	84	5.5	3.3	72	100.0%	0.80 [-0.57 , 2.17]
Subtotal (95% CI)			**84**			**72**	**100.0%**	**0.80 [-0.57 , 2.17]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.15 (P = 0.25)								
Total (95% CI)			**84**			**72**	**100.0%**	**0.80 [-0.57 , 2.17]**
Heterogeneity: Not applicable								
Test for overall effect: Z = 1.15 (P = 0.25)								
Test for subgroup differences: Not applicable								



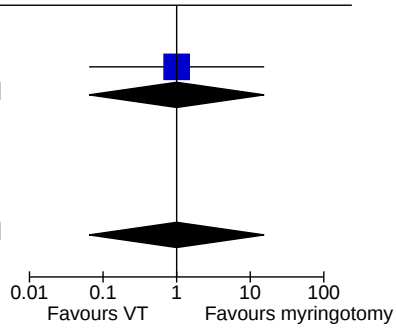
Footnotes

(1) Randomised by child, reported by ear at 12 months. Ave cluster size=2. DE=1

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 6: Sensitivity analysis. Mean final hearing threshold (medium-term). ICC=zero

Analysis 7.7

Study or Subgroup	Experimental Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Fixed, 95% CI
7.7.1 Sensitivity analysis: ICC = 1 (complete correlation between ears)						
Yousaf 2016 (1)	1	41	1	41	100.0%	1.00 [0.06 , 15.45]
Subtotal (95% CI)		**41**		**41**	**100.0%**	**1.00 [0.06 , 15.45]**
Total events:	1		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						
Total (95% CI)		**41**		**41**	**100.0%**	**1.00 [0.06 , 15.45]**
Total events:	1		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.00 (P = 1.00)						
Test for subgroup differences: Not applicable						



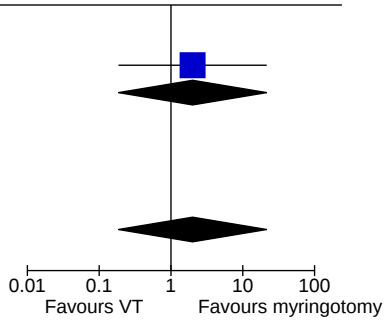
Footnotes

(1) VT versus laser myringotomy. Randomised by child, reported by ears. Ave cluster size=1.66. DE=1.66

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 7: Sensitivity analysis. Persistent perforation (medium-term). ICC=1.0

Analysis 7.8

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
7.8.1 Sensitivity analysis: ICC of zero (no correlation between ears)						
Yousaf 2016 (1)	2	68	1	68	100.0%	2.00 [0.19 , 21.54]
Subtotal (95% CI)		**68**		**68**	**100.0%**	**2.00 [0.19 , 21.54]**
Total events:	2		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.57 (P = 0.57)						
Total (95% CI)		**68**		**68**	**100.0%**	**2.00 [0.19 , 21.54]**
Total events:	2		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.57 (P = 0.57)						
Test for subgroup differences: Not applicable						



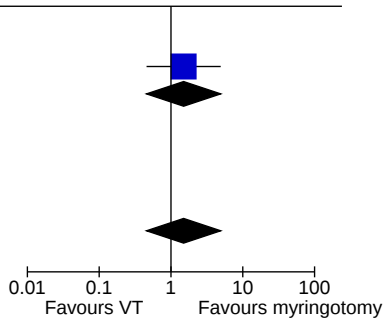
Footnotes

(1) VT versus laser myringotomy. Randomised by child, reported by ears. Ave cluster size=1.66. DE=1

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 8: Sensitivity analysis. Persistent perforation (medium-term). ICC=zero

Analysis 7.9

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
7.9.1 Sensitivity analysis: ICC of 1.0 (complete correlation between ears)						
Yousaf 2016 (1)	6	41	4	41	100.0%	1.50 [0.46 , 4.92]
Subtotal (95% CI)		**41**		**41**	**100.0%**	**1.50 [0.46 , 4.92]**
Total events:	6		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.67 (P = 0.50)						
Total (95% CI)		**41**		**41**	**100.0%**	**1.50 [0.46 , 4.92]**
Total events:	6		4			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.67 (P = 0.50)						
Test for subgroup differences: Not applicable						



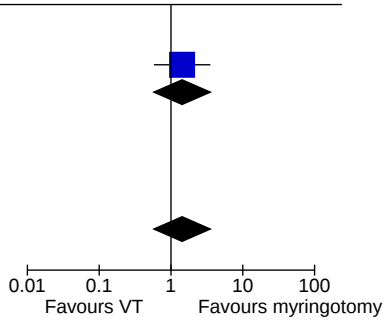
Footnotes

(1) VT versus laser myringotomy at 30 days. Reported by ear. Ave cluster size=1.66. DE=1.66

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 9: Sensitivity analysis. Persistence of OME: VT versus laser myringotomy (short-term). ICC=1.0

Analysis 7.10

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
7.10.1 Sensitivity analysis: ICC of zero (no correlation between ears)						
Yousaf 2016 (1)	10	68	7	68	100.0%	1.43 [0.58 , 3.53]
Subtotal (95% CI)		**68**		**68**	**100.0%**	**1.43 [0.58 , 3.53]**
Total events:	10		7			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.77 (P = 0.44)						
Total (95% CI)		**68**		**68**	**100.0%**	**1.43 [0.58 , 3.53]**
Total events:	10		7			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.77 (P = 0.44)						
Test for subgroup differences: Not applicable						



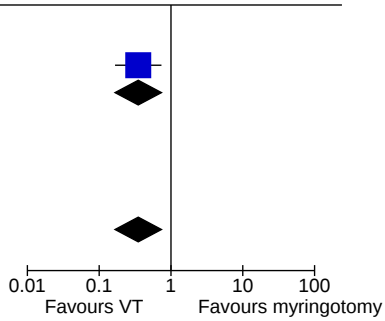
Footnotes

(1) VT versus laser myringotomy at 30 days. Reported by ear. Ave cluster size=1.66. DE=1

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 10: Sensitivity analysis. Persistence of OME: VT versus laser myringotomy (short-term) ICC=zero

Analysis 7.11

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
7.11.1 Sensitivity analysis: ICC of 1.0 (complete correlation between ears)						
Yousaf 2016 (1)	7	41	20	41	100.0%	0.35 [0.17 , 0.74]
Subtotal (95% CI)		**41**		**41**	**100.0%**	**0.35 [0.17 , 0.74]**
Total events:	7		20			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.77 (P = 0.006)						
Total (95% CI)		**41**		**41**	**100.0%**	**0.35 [0.17 , 0.74]**
Total events:	7		20			
Heterogeneity: Not applicable						
Test for overall effect: Z = 2.77 (P = 0.006)						
Test for subgroup differences: Not applicable						



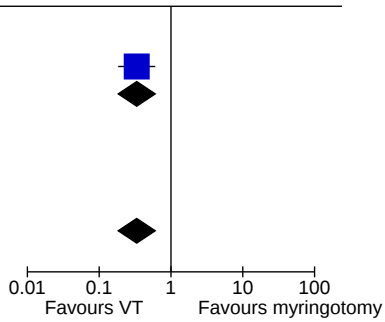
Footnotes

(1) At 6 months. Reported by ear. Ave cluster size=1.66. DE=1.66

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 11: Sensitivity analysis. Persistence of OME: VT versus laser myringotomy (medium-term). ICC=1.0

Analysis 7.12

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI
7.12.1 Sensitivity analysis: ICC of zero (no correlation between ears)						
Yousaf 2016 (1)	11	68	33	68	100.0%	0.33 [0.18 , 0.60]
Subtotal (95% CI)		**68**		**68**	**100.0%**	**0.33 [0.18 , 0.60]**
Total events:	11		33			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.63 (P = 0.0003)						
Total (95% CI)		**68**		**68**	**100.0%**	**0.33 [0.18 , 0.60]**
Total events:	11		33			
Heterogeneity: Not applicable						
Test for overall effect: Z = 3.63 (P = 0.0003)						
Test for subgroup differences: Not applicable						



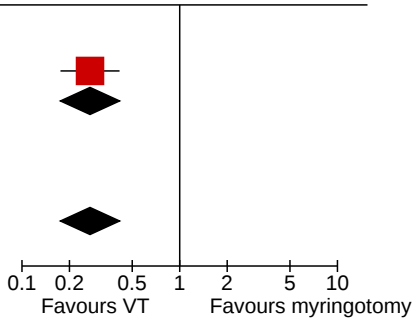
Footnotes

(1) At 6 months. Reported by ear. Ave cluster size=1.66. DE=1

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 12: Sensitivity analysis. Persistence of OME: VT versus laser myringotomy (medium-term). ICC=zero

Analysis 7.13

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI
7.13.1 Sensitivity analysis: correlation coefficient of 0.3 assumed				
Koopman 2004 (1)	-1.309333	0.22	100.0%	0.27 [0.18 , 0.42]
Subtotal (95% CI)			**100.0%**	**0.27 [0.18 , 0.42]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 5.95 (P < 0.00001)				
Total (95% CI)			**100.0%**	**0.27 [0.18 , 0.42]**
Heterogeneity: Not applicable				
Test for overall effect: Z = 5.95 (P < 0.00001)				
Test for subgroup differences: Not applicable				



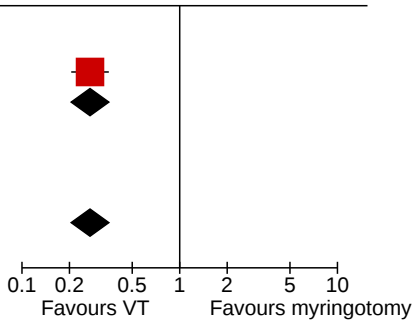
Footnotes

(1) Randomised by ear. Non-paired data. At 6 months.

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 13: Sensitivity analysis. Persistence of OME: VT versus laser myringotomy, randomised by ear (medium-term). CC=0.3

Analysis 7.14

Study or Subgroup	log[Odds Ratio]	SE	Weight	Odds Ratio IV, Random, 95% CI	Odds Ratio IV, Random, 95% CI
7.14.1 Sensitivity analysis: correlation coefficient of 0.7 assumed					
Koopman 2004 (1)	-1.309333	0.14	100.0%	0.27 [0.21 , 0.36]	
Subtotal (95% CI)			**100.0%**	**0.27 [0.21 , 0.36]**	
Heterogeneity: Not applicable					
Test for overall effect: Z = 9.35 (P < 0.00001)					
Total (95% CI)			**100.0%**	**0.27 [0.21 , 0.36]**	
Heterogeneity: Not applicable					
Test for overall effect: Z = 9.35 (P < 0.00001)					
Test for subgroup differences: Not applicable					



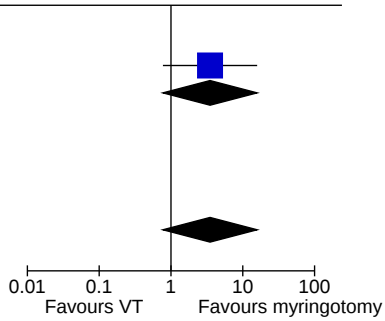
Footnotes

(1) Randomised by ear. Non-paired data. At 6 months.

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 14: Sensitivity analysis. Persistence of OME: VT versus laser myringotomy, randomised by ear (medium-term). CC=0.7

Analysis 7.15

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
7.15.1 Sensitivity analysis: ICC of 1.0 (complete correlation between ears)							
Yousaf 2016 (1)	7	41	2	41	100.0%	3.50 [0.77 , 15.85]	
Subtotal (95% CI)		**41**		**41**	**100.0%**	**3.50 [0.77 , 15.85]**	
Total events:	7		2				
Heterogeneity: Not applicable							
Test for overall effect: Z = 1.63 (P = 0.10)							
Total (95% CI)		**41**		**41**	**100.0%**	**3.50 [0.77 , 15.85]**	
Total events:	7		2				
Heterogeneity: Not applicable							
Test for overall effect: Z = 1.63 (P = 0.10)							
Test for subgroup differences: Not applicable							



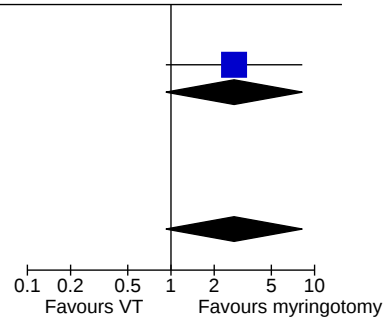
Footnotes

(1) At 6 months. Reported by ear. Ave cluster size=1.66. DE=1.66

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 15: Sensitivity analysis. Retraction of TM: VT versus laser myringotomy (medium-term). ICC=1.0

Analysis 7.16

Study or Subgroup	Experimental Events	Experimental Total	Control Events	Control Total	Weight	Risk Ratio M-H, Random, 95% CI	Risk Ratio M-H, Random, 95% CI
7.16.1 Sensitivity analysis: ICC of zero (no correlation between ears)							
Yousaf 2016 (1)	11	68	4	68	100.0%	2.75 [0.92 , 8.21]	
Subtotal (95% CI)		**68**		**68**	**100.0%**	**2.75 [0.92 , 8.21]**	
Total events:	11		4				
Heterogeneity: Not applicable							
Test for overall effect: Z = 1.81 (P = 0.07)							
Total (95% CI)		**68**		**68**	**100.0%**	**2.75 [0.92 , 8.21]**	
Total events:	11		4				
Heterogeneity: Not applicable							
Test for overall effect: Z = 1.81 (P = 0.07)							
Test for subgroup differences: Not applicable							



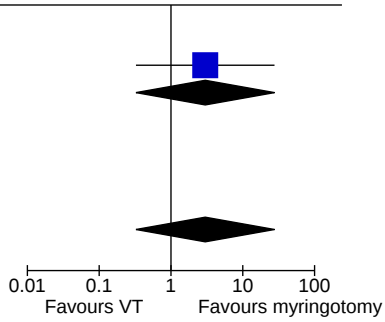
Footnotes

(1) At 6 months. Reported by ear. Ave cluster size=1.66. DE=1

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 16: Sensitivity analysis. Retraction of TM: VT versus laser myringotomy (medium-term). ICC=zero

Analysis 7.17

Study or Subgroup	Experimental Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
7.17.1 Sensitivity analysis: 1.0 (complete correlation between ears)						
Yousaf 2016 (1)	3	41	1	41	100.0%	3.00 [0.33 , 27.66]
Subtotal (95% CI)		**41**		**41**	**100.0%**	**3.00 [0.33 , 27.66]**
Total events:	3		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.97 (P = 0.33)						
Total (95% CI)		**41**		**41**	**100.0%**	**3.00 [0.33 , 27.66]**
Total events:	3		1			
Heterogeneity: Not applicable						
Test for overall effect: Z = 0.97 (P = 0.33)						
Test for subgroup differences: Not applicable						



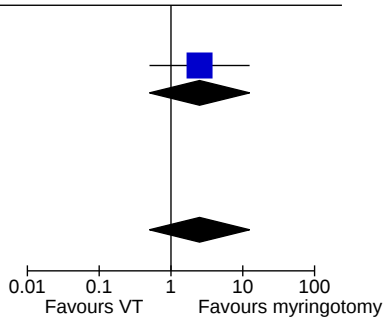
Footnotes

(1) At 6 months. Reported by ear. Ave cluster size=1.66. DE=1.66

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 17: Sensitivity analysis. Otorrhoea: VT versus laser myringotomy (medium-term). ICC=1.0

Analysis 7.18

Study or Subgroup	Experimental Events	Total	Control Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
7.18.1 Sensitivity analysis: ICC of zero (no correlation between ears)						
Yousaf 2016 (1)	5	68	2	68	100.0%	2.50 [0.50 , 12.44]
Subtotal (95% CI)		**68**		**68**	**100.0%**	**2.50 [0.50 , 12.44]**
Total events:	5		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.12 (P = 0.26)						
Total (95% CI)		**68**		**68**	**100.0%**	**2.50 [0.50 , 12.44]**
Total events:	5		2			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.12 (P = 0.26)						
Test for subgroup differences: Not applicable						



Footnotes

(1) At 6 months. Reported by ear. Ave cluster size=1.66. DE=1

Comparison 7: Sensitivity analyses: VT versus myringotomy, Outcome 18: Sensitivity analysis. Otorrhoea: VT versus laser myringotomy (medium-term). ICC=zero