Evidence Report/Technology Assessment

Number 55

Diagnosis, Natural History, and Late Effects of Otitis Media With Effusion

Volume 1. Evidence Report and Evidence Tables

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
2101 East Jefferson Street
Rockville, MD 20852
www.ahrq.gov

Contract No: 290-97-0001, Task Order No. 04

Prepared by:
Southern California Evidence-based Practice Center

EPC Director
Paul Shekelle, M.D., Ph.D.

Principle Investigators
Glenn Takata, M.D.
Linda S. Chan, Ph.D.

Investigators
Rita Mangione-Smith, M.D.
Pamela M. Corley, MSLS
Tricia Morphew, M.S.
Sally Morton, Ph.D.

AHRQ Publication No. 03-E023
May 2003

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

AHRQ is the lead Federal agency charged with supporting research designed to improve the quality of health care, reduce its cost, address patient safety and medical errors, and broaden access to essential services. AHRQ sponsors and conducts research that provides evidence-based information on health care outcomes; quality; and cost, use, and access. The information helps health care decisionmakers—patients and clinicians, health system leaders, and policymakers—make more informed decisions and improve the quality of health care services.

Suggested Citation:
Shekelle P, Takata G, Chan L, et al. Diagnosis, Natural History, and Late Effects of Otitis Media with Effusion. Evidence Report/Technology Assessment No. 55 (Prepared by Southern California Evidence-based Practice Center under Contract No 290-97-0001, Task Order No. 4). AHRQ Publication No. 03-E023. Rockville, MD: Agency for Healthcare Research and Quality. May 2003.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-Based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To bring the broadest range of experts into the development of evidence reports and health technology assessments, AHRQ encourages the EPCs to form partnerships and enter into collaborations with other medical and research organizations. The EPCs work with these partner organizations to ensure that the evidence reports and technology assessments they produce will become building blocks for health care quality improvement projects throughout the Nation. The reports undergo peer review prior to their release.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

We welcome written comments on this evidence report. They may be sent to: Acting Director, Center for Practice and Technology Assessment, Agency for Healthcare Research and Quality, 6010 Executive Blvd., Suite 300, Rockville, MD 20852.

Carolyn Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Acting Director, Center for Practice and Technology Assessment
Agency for Healthcare Research and Quality

The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services of a particular drug, device, test, treatment, or other clinical service.

Structured Abstract

Objectives. The purpose of this evidence-based report is to review the evidence on the natural history of otitis media with effusion (OME), the late effects of early life otitis media on hearing and speech and language development, and the operating characteristics of various methods of diagnosing OME. OME is defined as "fluid in the middle ear without signs or symptoms of ear infection." The evidence compiled in this report is intended to aid clinicians, health care provider organizations, and others to develop clinical practice guidelines or medical review criteria for OME. The report also identified areas for future research.

Search Strategy. The MEDLINE search strategy used both controlled vocabulary MeSH (Medical Subject Headings) terms and keywords to ensure that all relevant citations were retrieved. Search terms for otitis media with effusion were combined with search terms for natural history, speech and language development, hearing, and diagnosis. The otitis media module included otitis media, otitis media with effusion, suppurative otitis media, allergic otitis media, fluid ear, glue ear, middle ear effusion, mucoid otitis media, nonsuppurative otitis media, secretory otitis media, and serous otitis media. The natural history terms included natural course, natural history, placebo, placebos, resolution, self-limited, self-limiting, untreated, and a variety of terms for spontaneous resolution. The speech and language module included speech and language, speech and language disorders, child language, communication, communication disorders, language development and tests, voice, and voice disorders. The hearing module included hearing and hearing disorders and hearing aids and tests, as well as the text word hearing. The diagnosis module used diagnosis and diagnostic techniques and procedures, as well as the text words audiometry, diagnosis, diagnostic, otoscopy, tympanometry history, speech and language development, hearing, and diagnosis.

Selection Criteria. Excluded were studies on patients with immunodeficiencies, craniofacial anomalies (including cleft palate), primary mucosal disorders, or genetic conditions. Prospective cohort studies were included for questions that addressed natural history, speech, language, and hearing. Prospective diagnostic studies were used to evaluate the operating characteristics of diagnostic methods.

Data Collection and Analysis. Two physicians or one physician and one health services researcher independently screened all titles and/or abstracts for potential inclusion, evaluated the quality of the articles, and abstracted data from full-length articles onto pre-designed forms. The selection criteria included human studies that addressed a key question about OME in children. Excluded were case reports, editorials, letters, reviews, practice guidelines, and non-English language publications.

Main Results. We found that 22.5 to 42.7 percent of OME in children older than 3 years of age cumulatively resolves over a period of three months, depending on the definition of OME resolution. Based on a limited number of cohort studies, we found no evidence to support an impact of early life otits media, defined as a history of otitis media at less than 3 years of age, on expressive language, receptive language, or cognitive verbal intelligence at age older than 3 years. However, this evidence is insufficient to exclude the possibility that a clinically important effect does exist, therefore strong conclusions cannot be drawn about the effect of otitis media at

an early age on subsequent speech and language development. The generalizability of this finding on speech and language is suspect because the populations represented by the six cohorts utilized in the meta-analyses were primarily those of particular ethnic/racial origin. Moreover, the findings cannot be generalized to children with craniofacial defects, primary mucosal disorders, immunodeficiencies, genetic conditions, or pre-existing developmental disorders, and may not necessarily be generalized to children with persistent bilateral otitis media. Children with early life otitis media have a higher risk of conductive hearing loss, defined using a threshold greater than or equal to 20 dB at any frequency with or without treatment, at age 6 to 10 years than children without early life otitis media. The pooled relative risk of conductive hearing loss was 2.6 (95% CI: 1.6 to 4.2). We found insufficient data to assess early-life OM on permanent (or sensorineural) hearing loss. Among nine diagnostic methods, pneumatic otoscopy had the best apparent performance with a sensitivity of 93.8 percent (95% CI: 91.4%, 96.3%) and a specificity of 80.5 percent (95% CI: 75.1%, 86.0%). However, tester qualifications were reported inconsistently, and training was not specified.

Conclusions. Although these estimates must be viewed with great caution due to heterogeneity that arose from study design and documentation issues for which we could not adjust in our analysis, about 22.5 to 42.7 percent cumulatively resolved over a period of three months, depending on the definition of OME resolution. Our findings on the possible effects of early life otitis media on speech and language development are in general agreement with the 1994 Agency for Healthcare Research and Quality OME guideline conclusion that rigorous, methodologically sound research does not adequately support or refute the possible effect of otitis media on speech and language. We found that children with early life otitis media have a greater risk of conductive hearing loss at age 6 to10 years. In addition, we found that pneumatic otoscopy had the best operating characteristics among the nine alternatives examined, for diagnosing the presence of middle-ear effusion in OME at single points in time.

Considering the abundance of literature addressing otitis media, these findings concerning natural history, speech and language development, and hearing are very limited. Future research on the natural history of otitis media with effusion must focus on improving study quality. In particular, control of therapeutic intervention during the study and the distinction between OME persistence and recurrence needs to be addressed. For evaluation of long-term effects of early life otitis media on speech, language, or hearing, a coordinated approach that uses uniform definitions and considers the interactions of multiple risk factors, interventions, and outcome measures is recommended. Such an integrated approach is also important for the evaluation of diagnostic methods. Further, a systematic review of diagnostic studies that employ algorithms or aggregated scores may be useful.

Contents

Evidence Report

Evidence Tables

Figures

Tables

Agency for Healthcare Research and Quality

Evidence Report/Technology Assessment

Number 55

Diagnosis, Natural History, and Late Effects of Otitis Media with Effusion

Summary

Overview

This evidence-based report reviews the evidence on the natural history of otitis media with effusion (OME), the impact of otitis media on long-term speech and language development and on hearing, and the operating characteristics of various methods of diagnosing otitis media with effusion. OME is defined as "fluid in the middle ear without signs or symptoms of ear infection." The evidence compiled in this report is intended to aid clinicians, health care provider organizations, and others to develop clinical practice guidelines or medical review criteria for OME. The report will also identify areas for future research.

Reporting the Evidence

Based on degree of importance (including level of controversy) and feasibility of answering the question, the Technical Expert Panel limited the scope of this evidence report to four key questions:

1) The natural history of otitis media with effusion (OME)?
2) The long-term effects of early-life otitis media, defined as positive otitis media history at less than three years of age, on speech and language development?
3) The long-term effects of early-life otitis media on hearing?
4) The operating characteristics of various methods of diagnosing otitis media with effusion?

Methodology

A 12-member Technical Expert Panel that consisted of clinical experts, a consumer, and a representative of a managed care organization convened to:

- advise the project in the ranking of proposed key questions and influencing factors
- guide the development of the scope and definition of OME
- advise in development of the search strategy, and
- review and comment on the plan of analysis.

The Technical Expert Panel and project staff developed a literature search strategy. Project staff searched MEDLINE® (1966-January 2000), the Cochrane Library (through January 2000), and EMBASE (1980-January 2000). Additional articles were identified by review of reference lists in proceedings, published articles, reports, and guidelines.

The MEDLINE® search strategy used both controlled vocabulary MeSH® (Medical Subject Headings) terms and keywords to ensure that all relevant citations were retrieved. The strategy included search terms for otitis media with effusion combined with search terms for natural history, speech and language development, hearing, and diagnosis.

The otitis media module included *otitis media, otitis media with effusion, suppurative otitis media, allergic otitis media, fluid ear, glue ear, middle-ear effusion, mucoid otitis media, nonsuppurative otitis media, secretory otitis media, and serous otitis media.*

The natural history terms included *natural course, natural history, placebo, placebos, resolution, self-limited, self limiting, and untreated,* as well as a variety of terms for spontaneous resolution.

The speech and language module included *speech and language, speech and language disorders, child language, communication, communication*

disorders, language development and tests, voice, and voice disorders.

The hearing module included *hearing and hearing disorders, hearing aids and tests*, and the text word *hearing.*

The diagnosis module used *diagnosis and diagnostic techniques and procedures*, as well as the text words *audiometry, diagnosis, diagnostic, otoscopy, and tympanometry*.

Two physicians or one physician and one health services researcher independently screened all titles and/or abstracts for potential inclusion, evaluated the quality of the articles, and abstracted data from full-length articles onto pre-designed forms. The selection criteria included human studies that addressed a key question about OME in children. Excluded were case reports, editorials, letters, reviews, practice guidelines, non-English language publications, and studies on patients with immunodeficiency disorders or craniofacial anomalies, including cleft palate.

For the natural history question, we used only prospective cohort(s) studies on untreated subjects from which outcome data were abstractable for children up through age 12 years. For the speech and language and hearing questions, we used only prospective cohort studies that fulfilled the following criteria: the degree of OME was determined during the first three years of life, upper age limit was 22 years, the degree of OM was graded in some way, and the outcome was measured when the child was older than age three years. For the diagnostic methods question, we used only prospective studies on children up through 12 years of age that fulfilled four criteria: the diagnostic procedure of interest was performed within 24 hours of the reference standard, was not an algorithm or combination of multiple diagnostic procedures, used one of the acceptable reference standards specified in the scope, and produced abstractable data.

The first step of all analyses was to obtain a distribution of studies stratified by the population characteristics, type of outcome measures, and non-treatment factors. This step provided us with an overview of the emphasis of past research in this area and an opportunity to identify gaps and areas for future research.

In strata with more than three studies, we performed a meta-analysis for a pooled random effects estimate of an outcome with 95% confidence intervals. In addition to deriving the pooled estimate, we evaluated the heterogeneity of the study outcomes. For the evaluation of diagnostic methods, we estimated the sensitivity, specificity, and positive and negative predictive values for each diagnostic procedure compared to a particular reference standard with three or more studies.

This evidence report was reviewed by the Technical Expert Panel as well as an 18-member peer review panel that consisted of content experts, consumers, representatives of managed care organizations, an expert in pediatric pharmacology, and methodologists. All comments received from these individuals were reviewed and acted upon appropriately.

Findings

Natural History of OME

- No meta-analyses for children under three years of age were possible, because we could identify only two studies each for the under six months and the three-months-to-three-years age groups. For the over-three-years age group, two sets of meta-analyses showed that 22.5 to 42.7 percent of ears with OME cumulatively resolved over a period of three months, depending on the definition of OME resolution. These estimates must be viewed with caution due to the clinical heterogeneity evident in the data synthesized and due to the weaknesses of design or documentation of the study cohorts. In particular, in most cases investigators did not document whether subjects had received medical or surgical treatment during the course of the study that could affect OME outcome or how compliance with non-treatment was established. Of those investigators who reported how many children received treatment, the majority did not stratify their findings by treatment status.
- A few of the studies analyzed OME resolution by influencing factors such as gender, care at home versus daycare, season, side of affected ear, race or ethnicity, or diagnostic instrument. Because of the paucity of such studies, quantitative synthesis was not possible, and we refrain from drawing any conclusions regarding the effect of these influencing factors on resolution.
- As measured by scoring of documentation in the published articles, the quality of 28 cohort studies on natural history was generally poor.

- Half of the studies that attempted to study the natural history of OME did not control for or did not document control of interventions, either medical or surgical, that might affect OME outcome during the study period. The majority of these studies did not stratify findings by intervention status.
- The interval between examinations for OME in these studies ranged from one day to three years. For studies with long follow-up intervals, it was not possible to determine whether the presence of OME was due to persistence or recurrence. The criteria for follow-up varied among studies. Most studies continued follow-up for the duration of the study period regardless of the OME status at a particular exam, but four cohorts discontinued follow-up of individuals who had type A or normal tympanograms at any exam.

Early-life OM and Long-term Speech and Language Development

- Studies that addressed the effects of early-life otitis media on long-term speech and language development among children differed considerably with respect to risk factors studied, type of outcome measured, method of measurement, unit of measurement, age at outcome determination, and study design.
- The meta-analyses that could be conducted on long-term expressive language, receptive language, and cognitive verbal intelligence showed no effect of early otitis media as measured during the first three years of life. These findings may not be generalizable, since five of the six cohorts that were included in these three meta-analyses focused primarily on children from specific ethnic/racial groups or from particular socioeconomic groups. Furthermore, the results of these studies cannot be applied to children with craniofacial defects, primary mucosal disorders, immunodeficiency disorders, genetic conditions, or pre-existing developmental disorders, because children with these conditions were excluded from this analysis. In addition, only one of the studies included in these meta-analyses focused solely on persistent bilateral otitis media as opposed to unspecified unilateral or bilateral otitis media.

Early-life OM and Long-term Hearing

- Few studies on the effects of early-life otitis media on long-term hearing used a prospective cohort study design.
- Of the eight cohort studies analyzed, one set of four studies reported percentage of conductive hearing loss at six to ten years of age. For this analysis, the threshold for conductive hearing loss was defined as greater than or equal to 20 dB at any frequency, with or without treatment of otitis media.
- The pooled risk of conductive hearing loss at six to ten years among 346 children who had a positive history of early-life OM was 22 percent (95% CI: 7% to 36%). In contrast, the pooled risk of conductive hearing loss at six to ten years of age among 237 children with no history of early-life OM was 6 percent (95% CI: 1% to 12%). The pooled rate difference of conductive hearing loss at six to ten years of age between children with a positive OM history and those with a negative OM history was 11 percent (95% CI: 3% to 19%). Neither the studies pooled for the rate difference nor the studies pooled for the risk ratio showed significant heterogeneity in the outcomes.
- The findings were based on four homogeneous, though very different populations, one from Finland, another from Sweden, one primarily of American Indian children, and another primarily of Eskimo children. The four studies also differed on the definition and collection of OM history and on exclusion factors.
- We found insufficient data to assess the impact of early-life OM on permanent (sensorineural) hearing loss.

Diagnostic Methods for OME

- Based on our evaluation of 52 diagnostic studies, we were able to assess the ability of the following methods to diagnose middle-ear effusion in OME at a single point in time: acoustic reflectometry at ≤5 or >5 reflective units (RU); pneumatic otoscopy; portable tympanometry; professional tympanometry using acoustic reflex at 500 or 1000 Hz; professional tympanometry using static compensated acoustic admittance at 0.1, 0.2, and 0.3; professional tympanometry using B curve as abnormal; and professional tympanometry using B

or C2 curves as abnormal. All comparisons used myringotomy as the reference standard.

- Among the eight diagnostic methods, the receiver-operator characteristic points (plotting sensitivity against 1 minus specificity) showed that pneumatic otoscopy was closest to the optimal operating point where both sensitivity and specificity would be 100%. However, tester qualifications were reported inconsistently, and training was not specified. The pooled sensitivity was 94 percent (95% CI: 91%, 96%) and the pooled specificity was 80 percent (95% CI: 75%, 86%). These findings were based on 2,694 children from seven studies that reported a pooled prevalence of OME of 63 percent (95% CI: 58%, 67%). The prevalence rate ranged from 56 percent to 71 percent, which indicated significant heterogeneity among outcomes ($p<0.001$).

Limitations of the Literature

- **Natural History of OME:** Literature on the natural history of otitis media with effusion was difficult to interpret because of its generally poor quality, the lack of control for therapeutic interventions, the inability to distinguish persistent from recurrent OME due to the length of follow-up intervals, and the varied criteria for continued follow-up from exam to exam. Differing definitions of OME resolution and diagnostic methods made comparison difficult. Few studies considered the child or the episode as the unit of analysis, included younger children, or assessed types of OME other than newly diagnosed OME of unknown duration. In addition, few studies addressed the possible effects of influencing factors on OME resolution.
- **Early-life OM and Long-term Speech and Language Development:** The literature on the long-term effects of early-life otitis media on speech and language development diverged considerably with respect to methodology. As a result, findings could not be combined easily.
- **Early-life OM and Long-term Hearing:** Although the literature on the long-term effects of early-life otitis media on hearing was abundant, few studies used a prospective cohort study design. Because of the limited nature of this evidence and because the rate of intervention is highly dependent on the threshold hearing level adopted, the findings of this analysis should be applied with caution.
- **Diagnostic Methods for OME:** Nine comparisons of diagnostic methods enabled derivations of pooled estimates of diagnostic accuracy. However, more comparisons could not be made, including those that would have evaluated clinical signs and/or symptoms, air and/or bone threshold audiometry, binaural micro-tympanoscopy, and non-pneumatic otoscopy. Diagnostic methods that use algorithms or aggregated scorings are important but are not included in this evidence assessment.

Future Research

Future research on the natural history of otitis media with effusion must focus on improvement of study quality and establishing the effect of OME on long-term outcomes such as speech, language, and hearing. In particular, control of therapeutic interventions during the study and the distinction between OME persistence and recurrence need to be addressed. Adopting a less restrictive definition of non-intervention might simplify the analysis of studies of the natural history of OME. In addition, researchers, in conjunction with clinicians, should agree upon standard procedures for follow-up, including intervals of follow-up, definition of OME resolution, and diagnostic methods, so that resolution rates are indeed comparable. Future research must consider the child as the unit of analysis, since the outcomes of ultimate interest, such as speech, language, and hearing, are functional requirements of a child, not an ear. More research is needed on the role of influencing factors on the natural history of OME, so that the clinician on a particular day in a particular setting can make a better decision when assessing a particular child with particular characteristics.

Evaluation of long-term effects of early-life otitis media on speech, language, or hearing requires a coordinated systematic approach that uses a rational conceptual framework. Such an approach should address the risk factors, interventions, and outcome measures in an integrated fashion. The definition, classification, and type and unit of measure should be developed by a team of experts with the goal of standardizing definitions and approaches. Literature on findings should report both univariate and multivariate findings to enhance understanding of the patient and study characteristics and to allow pooling of data. An

integrated approach is also important for the evaluation of diagnostic methods. Such an approach will provide guidance for future studies. Future studies of diagnostic assessments of OME also should consider cost-effectiveness analysis, which can take into account the variable proficiency of clinicians in performing pneumatic otoscopy as well as the consequences of testing and patient preferences. Cost-effectiveness analysis will lead to a more informed decision on the best diagnostic method for OME.

Availability of Full Report

The full report from which this summary was taken was prepared for AHRQ by the Southern California Evidence-based Practice Center/RAND under contract No. 290-97-0001. It is expected to be available in summer 2002. When available, printed copies may be obtained free of charge from the AHRQ Publications Clearinghouse by calling 1-800-358-9295. Requesters should ask for Evidence Report/Technology Assessment No. 55, *Diagnosis, Natural History, and Late Effects of Otitis Media with Effusion* (AHRQ Publication No. 02-E026). Internet users will be able to access the report online through AHRQ's World Wide Web site www.ahrq.gov.

AHRQ Pub. No. 02-E025
June 2002
ISSN 1530-440X

Evidence Report

Chapter 1. Introduction

Purpose

The purpose of this evidence-based report is to review the evidence on the diagnosis, natural history, and late effects of otitis media with effusion (OME) on long-term speech, language, and hearing. The evidence compiled in this report is intended to aid clinicians, health care provider organizations, and others to develop clinical practice guidelines or medical review criteria for OME. The report will also identify areas for future research. Despite the relatively recent publication of the 1994 Otitis Media with Effusion in Children guideline (Stool, Berg, Berman, et al., 1994), the technical experts believed that OME remained a topic worthy of evidence-based inquiry due to the continued controversy over the care of children with OME and the potential availability of new information.

Scope of Work

The technical experts initially proposed 20 questions that addressed 10 broad areas related to the diagnosis and treatment of otitis media with effusion: (1) allergens, (2) natural history, (3) speech and language, (4) diagnostic methods, (5) surgical interventions, (6) hearing, (7) antibiotics, (8) steroids, (9) antihistamines and decongestants, and (10) alternative or complementary therapies. These twenty questions were ranked based on the following criteria: (1) degree of potential impact on OME outcomes and on future guideline development and (2) the feasibility of answering the question within the one year time frame and the availability of new information in the literature. The scope of this report covers the four highest ranked questions: (1) the natural history of otitis media with effusion (OME), (2) the long-term impact of early-life otitis media on speech and language, (3) the long-term impact of early-life otitis media on hearing, and (4) the accuracy of methods of diagnosis of otitis media with effusion.

Definition

The definition of otitis media has been a complicated issue. Ben H. Senturia, quoted in Bluestone (1999), stated that “In the past, there has been a confusion of terms, in part because of a failure to distinguish conceptually between the disease process, otitis media, and one of the manifestations of that disease process, namely otitis media with effusion. Otitis media is dynamic and at any one time should be considered a single point in a continuum of the disease process.”

Recent comments on the definition of OME point to some of the complex issues involved:

1. The OME guideline (Stool, Berg, Berman et al., 1994) defined OME as "fluid in the middle ear without signs or symptoms of ear infection." The guideline listed the following synonyms for OME: serous otitis media, secretory otitis media, allergic otitis media, catarrhal otitis media, nonsuppurative otitis media, mucoid otitis media, secondary otitis media, hydrotubotympanum, exudative catarrh, tubotympanitis, tympanic hydrops, glue ear, fluid ear, middle ear effusion, and tubotympanic catarrh.

2. Bluestone (1999) defined otitis media with effusion (OME) as "an inflammation of the middle ear with liquid collected in the middle-ear space. The signs and symptoms of acute infection are absent and there is no perforation of the tympanic membrane." He stated that middle-ear effusions can be acute (< 3 weeks), subacute (3 weeks to 3 months), or chronic (> 3 months). He also stated that researchers should precisely define OME.

3. Patterson and Paparella (1999) agreed that the different forms of otitis media (OM) are "interrelated and occur in a continuum." They recognized OME as one of the three major forms of OM, the other two being chronic OM (active or inactive) and silent OM. They classified OME as serous OM, purulent OM, and chronic OM. They divided serous OM into acute serous OM, chronic serous OM, and mucoid OM.

4. Jung and Hanson (1999) agreed that OM consists of various stages. Though they viewed purulent otitis media, serous otitis media, and mucoid otitis media as different stages, they considered OME to encompass all three, except the early stages of acute otitis media (AOM).

5. Paradise (1995) also agreed that "AOM and OME constitute elements in an otitis media disease spectrum, that there often is a transition zone between them and that the two conditions sometimes may be indistinguishable from each other diagnostically."

For this evidence report, the Technical Expert Panel decided to use the definition used in the OME Guideline (Stool, Berg, Berman et al., 1994): "fluid in the middle ear without signs or symptoms of ear infection."

Diagnosis

Various methods have been proposed for the diagnosis of OME. The OME guideline panel drew several conclusions regarding diagnosis of OME (Stool, Berg, Berman et al., 1994). They recommended the use of pneumatic otoscopy as the primary diagnostic method with tympanometry as a confirmatory diagnostic method. These recommendations were based on limited scientific evidence and strong panel consensus and on limited scientific evidence and expert opinion, respectively. The OME guideline panel found no evidence linking the outcome of algorithms that combine the results of pneumatic otoscopy and tympanometry to the presence of middle-ear effusion. In addition, the panel believed that the evidence was insufficient to make any recommendation regarding the use of acoustic reflectometry in the diagnosis of OME. Finally, the panel decided not to make a recommendation on the use of tuning fork tests in the diagnosis of OME due to the lack of adequate studies. The OME guideline panel did not present any meta-analyses on diagnostic methods.

Pneumatic otoscopy is performed with a handheld unit that consists of a light source, a magnifying lens, and a speculum. The otoscope allows visual inspection of the tympanic membrane as well as the external ear canal. With the speculum securely in place, the degree of movement of the tympanic membrane in response to pneumatic pressure may be observed. Decreased tympanic membrane mobility in response to pneumatic pressure is believed to be related to the presence of middle-ear effusion as found in OME. (Carlson and Stool, 1999)

Tympanometry is performed by inserting into the ear a probe that emits a tone and measures the amount of sound energy reflected from the tympanic membrane as a function of ear canal air pressure. The instrument may or may not be handheld. The output of tympanometric measurement may be qualitative, that is, tympanogram patterns, or quantitative, for example static admittance, equivalent ear volume, tympanometric width, tympanometric peak pressure, or acoustic reflex. The flat or type B tympanogram is believed to be associated with the presence of middle-ear effusion. The type A tympanogram is believed to indicate normal middle-ear status. The relationship of the type C tympanogram to middle-ear status is less clear. (Carlson and Stool, 1999; Nozza, 1996)

The acoustic middle-ear muscle reflex, either ipsilateral or contralateral, may also be measured by acoustic emittance instruments and represents the contraction of the stapedius and tensor tympani in response to sound stimulation. Its absence may be related to the presence of middle-ear effusion depending on the clinical situation (Nozza, 1996).

Acoustic reflectometry is performed using a handheld instrument that measures the response of the tympanic membrane to a frequency-sweep sound spectrum. The spectral gradient angle, which is a function of the frequency and amplitude, may be related to middle-ear effusion presence. (Carlson and Stool, 1999; Nozza, 1996)

Evoked otoacoustic emissions are a measure of ear canal sounds that are generated in the cochlea. These sounds have the potential for clinical application (Nozza, 1996).

Audiometry measures hearing acuity, using behavioral or non-behavioral methods, at various sound frequencies. It is known that children may have decreased hearing in the presence of middle-ear effusion (Carlson and Stool, 1999). Although no "universal agreement" appears to exist regarding the definition of hearing loss, a hearing threshold no worse than 15 decibels (dB) is considered normal in children, and 20 dB may be considered normal in older children (Madell, 1999)

Epidemiology: Prevalence and Incidence

Accurate estimates of the prevalence or incidence of OME were not found, since published population-based estimates are not available on the specific diagnosis of OME. Data on office visits reported from the National Ambulatory Medical Care Survey (NAMCS) provide the best indication of prevalence and incidence of the disease, although nonsuppurative, suppurative, and unspecified otitis media were all grouped into the term otitis media, and OME was not separated from AOM in the analyses (Schappert, 1992; Schappert, 1996; Woodwell, 1997a; Woodwell, 1997b; Woodwell and Schappert, 1995). Gates (1996), commenting on the NAMCS data, stated, “for children it is probably safe to presume that AOME [AOM with effusion, i.e. AOM] is the principal disorder noted in these surveys.” The OME Guideline panel of the Agency for Health Care Policy and Research [presently the Agency for Healthcare Research and Quality (AHRQ)] estimated that 25 percent to 35 percent of the NAMCS visits for otitis media were for OME (Stool, Berg, Berman et al., 1994a).

Schappert (1992) reported on the 1975 to 1990 NAMCS data. Visits by patients younger than 15 years of age constituted 70.6 percent of all office visits with the principal diagnosis of otitis media in 1975, 78.9 percent of all office visits with the principal diagnosis of otitis media in 1980, 81.9 percent of all office visits with the principal diagnosis of otitis media in 1985, and 80.5 percent of all office visits with the principal diagnosis of otitis media in 1990. From 1975 to 1990, the percent of office visits with otitis media as the principal diagnosis increased among those less than 15 years of age; from 7.3 percent to 17.4 percent for children under 2 years old,

from 10.4 percent to 18.1 percent for the 2-5 year olds, from 6.9 percent to 10.5 percent for the 6-10 year olds, and from 2.6 percent to 5.2 percent for the 11-14 year olds. The number of visits with a principal diagnosis of otitis media per 100 persons per year for the same time period (1975 to 1990) increased from 31.5 to 102.1 for children less than 2 years of age, 20.8 to 47.8 for those 2-5 years of age, 10.2 to 18.2 for those 6-10 years of age, and 3.3 to 8.0 for those 11-14 years of age.

Rosenfeld (1994) noted that about a quarter of OME cases are discovered incidentally during well-child examinations. About 60 percent of children would have OME by 2 years old and 80 percent before school entry. The Agency for Healthcare Research and Quality 1994 OME guideline reported that in one study of children 2 to 6 years old in group child care, 53 percent had at least one episode of OME during the first year of the study, 61 percent had at least one episode during the second year of the study, and 30 percent had recurrent OME (Stool, Berg, Berman et al., 1994).

NAMCS also stratified data by specific physician type. From 1975 to 1990, the percent of office visits with a principal diagnosis of otitis media increased for pediatricians from 8.1 percent to 14.3 percent, for general practitioners and family physicians from 1.3 percent to 3.5 percent, and for otolaryngologists from 12.8 percent to 20.2 percent.

Data for 1975 to 1990 were also stratified by age. In 1990, the number of visits with a principal diagnosis of otitis media per 100 persons per year among children younger than 2 years was 62.9 for pediatricians, 24.0 for general practitioners and family physicians, and 9.1 for otolaryngologists. In 1990, the number of visits with a principal diagnosis of otitis media per 100 persons per year among children 2 to 5 years old was 29.0 for pediatricians, 11.4 for general practitioners and family physicians, and 6.6 for otolaryngologists (Schappert, 1992).

The reports on the NAMCS data for 1993 to 1996 did not stratify by age (Schappert, 1996; Woodwell, 1997a; Woodwell, 1997b; Woodwell and Schappert, 1995). However, if the 1993 to 1996 data were similar to that in 1975 to 1990, it would be reasonable to conclude that the majority of these patients were younger than 15 years of age.

NAMCS office visit data for 1993 to 1996 generally support earlier data. Suppurative and unspecified otitis media was the third most frequently listed principal diagnosis in 1993, the sixth most frequent in 1994, the fourth most frequent in 1995, and the seventh most frequent in 1996 for ambulatory care visits to physician offices, hospital outpatient departments, and emergency departments (Schappert, 1996; Woodwell, 1997a; Woodwell, 1997b; Woodwell and Schappert, 1995). In 1996, visits for a principal diagnosis of otitis media and eustachian tube disorders occurred 82.8 percent of the time in physician offices, 5.3 percent in hospital outpatient departments, and 11.9 percent in emergency departments (Woodwell, 1997a).

The NAMCS also provided data on the duration of office visits for otitis media. The percent of visits for otitis media of duration 6-15 minutes increased between 1975 to 1990 from 64 percent to 78 percent and was associated with a decrease in visits less than six minutes from 24 percent to 13 percent (Schappert, 1992). In terms of surgical procedures, the rate of ambulatory surgery visits per 10,000 population for those younger than 15 years of age for otitis media and eustachian tube disorders was 86.9 in 1994 and 83.9 in 1995, based on 498,000 and 484,000 visits respectively (Hall and Lawrence, 1997; Kozak, Hall, Pokras et al., 1997). In 1995, the number of myringotomy with tympanostomy tube placements reported by NAMCS was 521,000 for a rate of 90.2 such procedures per 10,000 children less than 15 years of age (Hall and Lawrence, 1997).

In general, the NAMCS data demonstrated the significance of otitis media—and by implication OME—based on the prevalence and incidence of the disease and the frequency and duration of visits and surgical interventions.

Burden of Illness Due to Otitis Media with Effusion

The treatment, complications and sequelae, and adverse effects of otitis media, including OME, are a substantial financial burden to the nation. Three estimates of the cost of otitis media (OM) are found in the literature (Gates, 1996b; Stool and Field, 1989; Stool, Berg, Berman et al., 1994). A fourth study reported the cost per episode of "persistent middle ear infection" (Berman, Roark, and Luckey, 1994). We have assessed the strengths and weaknesses of these four studies and presented our own estimate of otitis media costs in a previous evidence-based analysis (Takata, Chan, Shekelle, et al., in press).

Gates (1996b) placed the annual national total cost of acute otitis media at $3.15 billion for the 0- to 4-year old age group in an unspecified date in the early or mid-1990's and placed the cost of treatment of chronic otitis media with effusion at $1.854 billion dollars per year (Gates, 1996). Stool and Field (1989) put the national cost of otitis media at $2.4-3.4 billion in the 0- to 6-year old age group in an unspecified year presumably in the middle or late 1980's. Both Gates (1996b) and Stool and Field (1989) assumed prevalences of otitis media that were at variance with the only available national data on the utilization of care visits for otitis media. Stool, Berg, Berman, et al. (1994) presented the only estimate of the national cost of otitis media using a data-based estimate of costs per case. Using claims data from more than 100 health insurers, they estimated the overall average cost of treating a 2-year old child with OME in 1991 to be $1,330 and the national total cost of OME in 1991 to be $1.09 billion. Berman, Roark, and Luckey (1994) estimated the cost of treating persistent middle ear effusion in a 13-month old boy at $720-$1,372 under Colorado Medicaid reimbursement levels in 1992 and $1,265-$2,588 under private practice reimbursement rates. Not all these studies included the indirect cost of family caregiving services required due to a child having otitis media.

An attempt to provide an updated estimate on the cost of otitis media included insights into the cost of OME and chronic middle ear infection (Takata, Chan, Shekelle, et al., in press). This estimate was derived for the year 1995, referred to children under 18 years of age, and was based on reports of national, rather than regional, utilization for otitis media such as the National Ambulatory Medical Care Survey and the National Health Interview Survey. Using these national data sets, it was estimated that 2.22 million episodes of OME or chronic middle ear infection occurred in 1995. It was assumed that eighty percent of these episodes were unrelated to acute otitis media. Data from three sources (Stool, Berg, Berman et al., 1994; Berman, Rourke, and Lucky, 1994; and the U.S. Bureau of the Census, 1992,1996) were used to estimate direct, indirect, and total costs of $1,321, $490.25, and $1,811 for treatment of an episode of OME or chronic middle ear infection. For 1995, the total national cost of treating OME or chronic middle ear infection would stand at $4.02 billion.

Whether based on the four prior estimates of otitis media or OME cost or the more recent estimate, the economic burden of OME on the nation is large. Any effort to improve cost-effective care of OME will result in significant savings in national medical expenditures as well as improved quality of care provided to children with OME.

Chapter 2. Methodology

Nomination of Technical Experts

Eleven organizations were contacted for technical expert and peer reviewer nominations. They included the American Academy of Family Physicians (AAFP), the American Academy of Pediatrics (AAP), the American Academy of Otolaryngology-Head and Neck Surgery Foundation (AAO-HNS), the Ambulatory Pediatric Association (APA), the American Academy of Audiology (AAA), the American Speech-Language-Hearing Association (ASHA), the Society for Ear Nose and Throat Advances in Children (SENTAC), the National Association of Pediatric Nurse Associates and Practitioners (NAPNAP), the American Association of Health Plans (AAHP), the Foundation for Accountability (FACCT), and Family Voices**.**

Upon receiving nominations from the agencies, we identified 12 technical experts to serve on the panel. Included were two family physicians, two otolaryngologists, three pediatricians, one audiologist, one speech and hearing expert, one managed-care representative, one nurse practitioner, and one consumer. **Table 1** lists the membership of the Technical Expert Panel.

Topic Assessment and Refinement

A draft work plan for the topic assessment and refinement phase was mailed to the technical experts and representatives of our partners (the American Academy of Pediatrics, the American Academy of Family Practice, and the American Academy of Otolaryngology-Head and Neck Surgery) for review and comments together with a preliminary review that provided a summary of:

- Incidence and prevalence of otitis media with effusion, treatment and management alternatives, the characteristics and size of the affected populations, and the most affected practice settings and providers;

- The burden of illness associated with otitis media with effusion, including morbidity and mortality.

- Extent to which variation exists in practices associated with the prevention, diagnosis, treatment, or diagnosis and treatment of otitis media with effusion.

We reviewed and compiled four previously conducted evidence-based analyses on otitis media with effusion. Particular attention was given to the report of The Otitis Media Guideline Panel on "Managing Otitis Media with Effusion in Young Children" (Stool, Berg, Berman et al., 1994) We distributed these reports to the panel of technical experts for review and for preparation of the initial telephone conference call during which assessment and refinement of the topics wereassessed and refined.

Identification of Key Questions

Prior to the first conference call, we asked each technical expert to submit questions for consideration in the evidence report. The project team organized the responses and compiled an initial list of 20 key questions from the original task order and the letters from the nominating agencies (**Table 2**). This document was distributed to the technical experts after the first conference call, together with a polling form requesting the ranking of the top 10 key questions. The technical experts were asked to rank 10 of the 20 questions from 10 to 1, using 10 as the most important and 1 as the least important. The criteria for ranking were:

1) importance, which included
   a) potential impact on OME outcomes and
   b) b) potential impact on development of future OME guidelines by the partner organizations; and

2) feasibility, which included
   a) possibility of conducting a literature search, review, and data synthesis in 6 months,
   b) availability of sufficient information (data) in the literature to answer the question, and
   c) if applicable, sufficient new information (data) available to affect the results of the last systematic review of the question significantly.

The polling results are tabulated in **Table 3** and the comments are included in **Table 4**. The results were distributed to the experts for discussion during the second conference call.

The four top ranking key questions were selected for consideration in the evidence report. After several revisions at the suggestions of the technical experts, the wordings of the four top ranking (key) questions are as follows:

Key question 1: On Natural History

What is the natural history (spontaneous resolution rate over time without treatment) for:

- OME persisting after a discrete episode of acute otitis media,
- newly diagnosed OME of unknown duration (unilateral or bilateral),
- OME persisting for weeks or months (unilateral or bilateral),
- unilateral OME lasting 3 months or longer,
- bilateral OME lasting 3 months or longer?

Key question 2: On Speech and Language Development

a) Do infants and preschool children with longer duration early-life OME have greater delays in speech and language development (receptive or expressive) later in life as compared to those with shorter duration OME? One specific formulation of this question

is: Is OME-associated conductive hearing loss in the first 3 years of life a risk factor for speech and language developmental delays?

b) What are the risk factors that modulate the effect of OME on speech and language development in infants and preschool children?

Key question 3: On Hearing Decrease

a) Do infants and preschool children with longer duration early-life OME as compared to those with shorter duration OME have permanent (or sensorineural) hearing loss later in life? One specific formulation of this question is: Is OME-associated conductive hearing loss in the first 3 years of life a risk factor for permanent (or sensorineural) hearing loss later in life?

b) What are the risk factors that interact with the effect of OME on hearing later in life (unilateral or bilateral) in infants and preschool children?

Key question 4: On Diagnostic Methods

What are the sensitivity, specificity, and predictive values for alternative methods of diagnosing OME compared with one of the reference standards?

These methods include, but are not limited to:

- signs/symptoms
- non-pneumatic otoscopy
- pneumatic otoscopy, validated or un-validated examiner
- binocular micro-tympanoscopy
- portable tympanometer
- professional tympanometer
- quantitative tympanometry
- acoustic reflectometry (specify model and year)
- otoacoustic emissions
- audiometry, air or. bone conduction thresholds.

The reference standards to be used in evaluating these diagnostic tests will include tympanocentesis, sedated or non-sedated; MRI; myringotomy, sedated or non-sedated; validated pneumatic otoscopy; and CT Scan.

Identification and Refinement of Causal Pathways, Study Populations, Practice Settings, and Target Audience

The project staff developed a causal pathway and a scope for each of the key questions. We distributed these documents to the panel of technical experts for review and for preparation for the initial telephone conference call. During this call, we assessed and refined the topics and discussed the proposed key questions, target condition, patient populations, clinical context, interventions, and outcomes of interest. The following characteristics of outcomes were proposed: 1) outcomes would be divided into short term and long term; 2) long term outcomes would consist of percent time with effusion, frequency of acute otitis media, hearing loss, speech and language performance, cognition, academic achievement, and other developmental outcomes; 3) duration of short term outcomes would be defined as four weeks or less or eight weeks or less; and 4) duration of long term outcomes would be defined as greater than one year.

During this first conference call, the Technical Expert Panel decided to use the OME Guideline definition of OME: "fluid in the middle ear without signs or symptoms of ear infection". (Stool, Berg, Berman et al., 1994) Though the technical experts agreed on the definition of OME they could not agree on which signs or symptoms should be absent, i.e. what signs or symptoms differentiated OME from acute otitis media. During the first conference call, the technical experts advised us to avoid the use of the terms 'acute', 'subacute', or 'chronic' as descriptors of OME. Instead we should use the duration of OME, such as "under 3 months" versus "greater than or equal to 3 months" as descriptors.

Before the second conference call, the project team developed a draft of the conceptual framework for the proposed key questions. During the second conference call, the framework was discussed and the inputs of the experts were incorporated into the revised overall causal pathway (**Table 5**) from which the causal pathways of the final four key questions were developed. The causal pathways were distributed to the technical experts for further review and comment. The final version causal pathways for the four key questions are presented in **Tables 6 through 9**.

Based on the causal pathways and the discussions during the two conference calls, we developed the scope for each of the four key questions. The scope specifically defined the disease entity, study population, practice settings (including provider type), time period in practice setting, exclusion factors, interventions, influencing factors, outcome measures, literature sources, language, and study design for each key question to be included in the evidence report. We conducted a second poll of the technical experts on each of these domains in which we sought their approval, disapproval, or recommendations for revision on each domain. **Appendix B** presents the version of the scope distributed to the technical experts for polling of their comments and approval. **Appendix C** presents the results of the polling of the experts' comments on the scope. In response to these comments, we further revised the key questions, causal pathways, and the various domains of the scope. We incorporated comments from technical experts to the extent possible, except those related to other domains, those that were obvious misunderstandings or misinterpretations, or those suggestions for deletions or additions that could be handled during the analysis phase of the project. Specifically, the project

team retained several influencing factors recommended for deletion by some experts. The project team took note of the deletions and would later stratify the analysis, if possible, by the group of factors unanimously recommended by the experts versus those that were not. Such stratified analysis would depend on the number of studies that specifically address these factors. We revised the key questions according to experts' suggestions. We further revised the scope and reworded the key questions according to the final round of comments: the final version of the scope is included in **Appendix D**.

In preparation for supplemental analysis, we took polls to solicit the technical experts' opinion on the importance of the risk factors identified in the analytical framework. Specifically, we asked the experts the following questions for each factor:

a) Regarding Key Question 1: "Does this factor influence the natural history of OME?"

b) Regarding Key Question 2: "Does this factor have an independent effect on speech and language development separate from its effects on OME or unspecified OM?"

c) Regarding Key Question 3: "Does this factor have an independent effect on long-term hearing separate from its effects on OME or unspecified OM?"

d) Regarding Key Question 4: "Does this factor have an independent effect on the accuracy of a diagnostic method separate from its effects on OME or unspecified OM?"

The experts had a choice of responding "yes", "no", or "don't know." For each such opinion, the experts were asked to indicate the basis of their opinion by choosing one or more of the following: "Judgment/Experience," "Theoretical Construct," or "Literature".

The questionnaire for the two polls is included in **Appendix E** and the responses of the 12 technical experts are presented in **Appendix F.** A summary of the risk factors ranked by the importance assigned by the technical experts is presented in **Table 10**.

Literature Search

The Technical Expert Panel and project staff developed a literature search strategy. The literature search included the search of three databases: MEDLINE (1966-January2000), the Cochrane Library (through January 2000), and EMBASE (1980-January 2000). We identified additional articles by review of reference lists in proceedings, published articles, reports, and guidelines.

The project librarian developed an overall search strategy for MEDLINE (**Appendix G**) that incorporated the input from the technical experts and followed the scope of the project. The MEDLINE database is produced by the U.S. National Library of Medicine and is widely recognized as the premier source for bibliographic coverage of biomedical literature. It encompasses information from Index Medicus, Index to Dental Literature, and International Nursing, as well as other sources of coverage in the areas of allied health, biological and physical sciences, humanities, and information science as they relate to medicine and health care, communication disorders, population biology, and reproductive biology. We searched the MEDLINE database for publications dating back to 1966. Further, we included articles in the English language only for the following reasons. First, our experience with our evidence assessment of the management of acute otitis media (Takata, Chan, Shekelle et al., in press;

Chan, Takata, Shekelle et al., in press) demonstrated a low yield from non-English language publications (only two studies accepted out of 97 reviewed and both of theses were also published in English). Second, we needed to balance limited resources between reviewing non-English language literature and answering additional key questions. Since empiric evidence of the need to include non-English language literature in meta-analyses was mixed (Moher, Pham, Klassen et al., 2000) and reviewing non-English literature would be resource intensive, we chose to limit our scope to English language literature only.

The MEDLINE search strategy used both controlled vocabulary terms and keywords. The strategy was organized into modules or clusters of search statements. The main groupings included: otitis media with effusion (OME), mastoid, otitis media; natural history; speech and language; hearing; and diagnosis. These groupings corresponded to the key questions.

For the "otitis media with effusion" concept, both the controlled vocabulary term otitis media with effusion and text word were used. A variety of additional terms were used; such as allergic otitis media, fluid ear, glue ear, middle ear effusion, mucoid otitis media, nonsuppurative otitis media, secretory otitis media, and serous otitis media. For the "mastoid" concept, both the controlled vocabulary and the text word were used. The otitis media module included what is referred to as an "explode" of otitis media, which included the controlled vocabulary headings "otitis media", "otitis media with effusion", and "otitis media, suppurative."

The "natural history" module combined "OME" or "mastoid" with a combination of text words and controlled vocabulary terms for natural history including "natural course", "natural history", "placebo", "placebos", "resolution", "self limited", "self limiting", "untreated", and a variety of terms for spontaneous resolution.

Both the "speech and language" module and the "hearing" module combined OME, or mastoid, or an explode of "otitis media" with the speech, language, and hearing concepts. The speech and language component used the controlled vocabulary terms for speech and language, speech and language disorders, child language, communication, communication disorders, language development and tests, voice, and voice disorders. In addition, the text words "speech" and "language" were added. The hearing module used the controlled vocabulary terms for hearing and hearing disorders and hearing aids and tests, as well as the text word "hearing".

The "diagnosis" module combined OME or mastoid with a combination of text words and controlled vocabulary terms for diagnosis. In addition to the controlled vocabulary terms for diagnosis and diagnostic techniques and procedures, a number of text words were added for audiometry, diagnosis, diagnostic, otoscopy, and tympanometry.

We customized the search strategy initially developed for MEDLINE for EMBASE. EMBASE, the Excerpta Medica database, produced by Elsevier Science, is a major biomedical and pharmaceutical database that indexes over 3,800 international journals. EMBASE is one of the most widely used biomedical and pharmaceutical databases. The database currently contains over 6 million records, with more than 400,000 citations and abstracts added yearly. We searched the EMBASE database for citations dating back to 1980. For the search in the Cochrane Library, we used "otitis media" as the search term. The Cochrane Library contains several databases: (1) The Cochrane Database of Systematic Reviews, which contains Cochrane reviews published by the Cochrane Collaboration, an international organization dedicated to applying evidence-based-medicine principles to the review of important clinical topics; (2) The Cochrane Controlled Trials Register which is a bibliographic database of controlled trials; (3) The Database of Abstracts of Reviews of Effectiveness (DARE), which includes structured abstracts of systematic reviews that have been critically appraised by reviewers at the NHS Centre for Reviews and Dissemination in York and by other experts, for example from the American College of Physicians' Journal Club and the journal Evidence-Based Medicine; and (4)

The Cochrane Review Methodology Database which is a bibliography of articles on the science of research synthesis.

The Cochrane Library search yielded 666 titles/abstracts. The MEDLINE search resulted in 2,379 titles/abstracts. After eliminating duplicates, we retained 2,207 titles/abstracts for screening. The EMBASE search retrieved 1980 citations. After eliminating duplicates, we retained 327 for screening.

We conducted all searches in January 2000 and subjected a total of 3,200 titles/abstracts to screening by two physician reviewers. By merging and eliminating duplicates from the titles/abstracts from the three databases, we created a database of titles and abstracts.

EndNote software (EndNote Windows Version 3.0, 1st Edition. Niles Software Inc., Berkeley, CA) was used to keep a complete record of all titles/abstracts and identify duplications. This software stores, organizes, and tracks references by source (e.g. identified in MEDLINE), search strategy (date of search, index code specifying search criteria used), and a unique identification (UI) code for each article (assigned by the source used to find article). Electronic removal of duplicate citations was supplemented by manual cross-checking. In the event an article was identified by an expert panel member or through reference checking, the title and author of the reference were entered into MEDLINE through the Ovid search system (Ovid Technologies, Inc. 1998, Version: 7.8 Millennium source ID 1.3932.1.156.1.7, Revision: 1.303.2.8) to determine the UI. If a UI could not be found for the article, an alternate identification code was assigned.

EndNote assigned a record number to each new reference added to the master file. This number would not change once an article was added to the list and was used, in addition to the UI, to sort references for article retrieval and review.

Upon completing the literature search and duplicate checking, we exported the master list generated from EndNote to a Microsoft Excel spreadsheet for data export and analysis. We added codes including status of article retrieval, reviewer, and the results of the review.

Review of Retrieved Titles/Abstracts Against Screening Criteria

After retrieving of titles and abstracts from the literature search, two physician reviewers reviewed the abstracts against the inclusion/exclusion criteria to determine eligibility for inclusion in the evidence synthesis as defined in the scope and key questions. Titles/abstracts were not masked prior to review. A pre-designed screening form was used to record the reviews. A meeting was held to review the instructions for screening (**Appendix H**), including the use of the computerized data forms. The reviewers entered the screening results directly into the computer and forwarded the results electronically to our data analyst for processing. The screening results for each title/abstract were matched between the two reviewers, and discrepancies on inclusion or exclusion were resolved in conference calls among the two reviewers and the task order coordinator. The data analyst generated summary reports indicating those abstracts that passed the screening criteria and those that failed and the reasons for failure.

We completed screening of the 3,200 titles/abstracts from the Cochrane Library, MEDLINE, and EMBASE. After resolution of 376 discrepant citations, 2230 (70%) were rejected and 970 were accepted for full article review. The reasons for rejection of the 2230 citations are presented in **Table 11**. We also screened the database provided to us by the American Academy of Pediatrics (AAP) from its recent review of the topic. Of a total of 1918 titles/abstracts screened against our database, we identified 477 duplicates from our ENDNOTE database, leaving 1441 records from the AAP files that required further screening by the two reviewers.

The screening of the 1441 citations from AAP files identified 32 additional citations that required full article review.

The third source of reference material was the six proceedings of the International Symposium on Recent Advances in Otitis Media with Effusion from which we identified 159 additional citations for full article review.

The fourth source consisted of references in books and articles from which we identified 31 full articles for further review.

Retrieval and Review of Full Articles

The titles/abstracts identified for further review were forwarded to the library for full article retrieval. Libraries at both the Los Angeles County - University of Southern California Medical Center and the University of Southern California Health Sciences Campus were the primary sources of the articles. Those not found were retrieved through the Inter-Library Loan Program.

Because a large number of titles/abstracts had inadequate information for full evaluation, a secondary screening of full articles was conducted. Two physicians or a physician and a health services researcher reviewed each article. Articles were not masked prior to review. Discrepancies on inclusion/exclusion were resolved between the reviewers.

Of the 1,250 full length articles to be retrieved, 3 were irretrievable due to incorrect citation information. Secondary review of the remaining 1,247 full length articles from the various sources resulted in the rejection of 798 articles and acceptance of 449 articles: 141 for question 1, 112 for question 2, 186 for question 3 and 75 for question 4. **Table 12** provides the reasons for rejection of the 798 articles and **Table 13** summarizes the number of articles accepted during the secondary review process. During the fourth conference call, the experts raised the age limit to 22 for the responses to questions 2 and 3 to allow for detection of speech, language, and hearing problems past age 12, the original upper age limit. As a result, we revisited all titles/abstracts and articles that had been rejected because of the age limit, and four previously rejected articles were accepted from our original databases.

After establishing the analytical plan and before data abstraction, a physician and one health services researcher carried out a tertiary review of the 449 articles according to a set of established criteria for each key question. During this tertiary review, study design and quality were also evaluated.

For Question 1, we used three criteria for tertiary screening: 1) was the study a prospective cohort(s) study or a randomized control trial (RCT), 2) whether the control group in the RCT used the other ear as the 'unit of control' or not, and 3) were the outcome data abstractable?

For Questions 2 and 3, we established 5 criteria for tertiary screening:

1) Was the degree of OME determined for the first 3 years of life, and could the OM degree for the period before 3 years of age be linked to a specific outcome? If a study began at age 3, the study was not considered to fulfill this criterion.

2) Was the upper age limit 22 years of age? If a study included subjects older than 22 years, this criterion could be fulfilled only if outcomes for the 22 years of age and under was reported.

3) Was the degree of OM graded in some way, such as total time with OM, some measure of OM persistence, OM recurrence, or some measure of OM severity, and could the OM degree grade be linked to a specific outcome?

4) Was the study prospective? A study is considered prospective if the outcomes were measured prospectively. Cross-sectional and case-control studies were specifically excluded. A study that followed subjects prospectively for both OM history and outcome measures was considered a prospective cohort study. Studies that followed subjects prospectively for outcome measures (i.e. over a period of time) but retrospectively for OM history were considered to be retrospective-prospective studies and were accepted for inclusion. Studies that collected data on the outcome measures at one point in time, were considered retrospective cohort studies regardless of whether OM history was collected prospectively or retrospectively and were excluded. Studies that presented a cross-sectional analysis of prospectively collected data were not considered prospective and were excluded. Randomized controlled studies of an intervention with longitudinally measured outcomes were considered prospective and were included.

5) Was the outcome measured when the child was older than 3 years of age?

For Question 4, we established 3 criteria for inclusion during tertiary review: 1) Were the diagnostic procedure of interest and the reference standard performed within 24 hours of each other? 2) Is the diagnostic procedure not an algorithm or combination of multiple diagnostic procedures, 3) Are the reference standards one of those specified in the scope (tympanocentesis, MRI, myringotomy, validated pneumatic otoscopy, or CT scan), and 4) Are the data abstractable.

After tertiary review and data abstraction, a total of 114 articles were included in this evidence report, five of which addressed more than one key question:

- 38 cohort studies for question 1, the natural history question;
- 21 studies for question 2, the speech and language question;
- 8 studies for question 3, the hearing question; and
- 52 studies for question 4, the diagnostic method question.

Table 14 presents the results of tertiary review of the 449 articles.

Review and Assessment of Study Quality

We established criteria used for the assessment of study quality prior to the review of articles. Only prospective cohort studies were reviewed for Questions 1, 2 and 3 because of concerns about the validity of case-control, cross-sectional, and retrospective cohort studies. Diagnostic studies were reviewed for Question 4. The criteria used to evaluate the quality of

both types of studies were modified from the work by the McMaster University Group (Jaeschke, Guyatt, and Sackett, 1994; Sackett, 1981; Trout, 1981; Tugwell, 1981).

The quality of natural history studies was evaluated against the following criteria:

1) Was the study a prospective cohort study?

2) Was the outcome(s) of the study clearly defined?

3) Was the outcome(s) measured at a clearly defined timepoint(s)?

4) Was the cohort of subjects followed without any intervention?

6) Was there blinded assessment of the outcome(s) of the study?

7) Were point estimates and measures of variability provided for the main adverse outcomes measured?

The quality of prospective cohort studies was evaluated against eight components:

1) Was the study cohort(s) clearly defined, with clearly spelled out inclusion and exclusion criteria?

2) Was the study cohort(s) assembled at a uniform point in the course of the child's illness?

3) Were the pathways by which patients entered the study clearly described?

4) Was complete follow-up achieved?

5) Were withdrawals and drop-outs described?

6) Were objective outcome criteria developed and used?

7) Was the outcome assessment "blind"?

8) Was adjustment for extraneous factors carried out?

The quality of diagnostic studies was evaluated against six components:

1) Was the reference standard appropriate?
2) Were the test results and the reference standard assessed independently of each other?

3) Were the readers of the results of the diagnostic test or the reference standard blinded?

4) Did the patient sample include an appropriate spectrum of mild and severe, treated and untreated patients to whom the diagnostic tests were applied in clinical practice?

5) Were the reproducibility of the test result (precision) and its interpretation (observer variation) determined?

6) Were the methods for performing the test described in sufficient detail to permit replication?

Articles were not masked prior to review. The Task Order Coordinator resolved minor discrepancies between the two reviews of each article. Conferences were held to resolve discrepancies whenever needed.

Data Abstraction

For the articles deemed eligible for inclusion in the Evidence Report, data abstraction was carried out by a two-member team that consisted of a physician reviewer and a health services researcher. One of the two members abstracted the data onto the evidence table, and the other member checked the data for accuracy. Data abstracted included parameters necessary to define study groups, inclusion/exclusion criteria, influencing factors, and outcome measures to be used in analysis.

Specific instructions for data abstraction were recorded. For Question 1, the outcome indicators for abstraction included partial OME resolution (resolution in one ear for bilateral OME only), complete OME resolution, relapse/recurrence (fluctuation/dynamic course), AOM after OME. For Question 2, the outcome indicators for abstraction were expressive or receptive language, expressive or receptive speech, and cognitive verbal intelligence. For Question 3, the outcome indicators for abstraction included conductive or sensorineural hearing loss. For Question 4, the outcome indicators for abstraction were sensitivity, specificity, positive and negative predictive values, and prevalence rate. For all questions, the time or age at which each outcome was measured was recorded. The outcome measures included both continuous and categorical measures. Continuous measurements included mean time or a median time and the categorical measures included proportion with resolution at specified times where both numerators and denominators were recorded. A key issue was that individual children, not populations, must be tracked. The latter would be acceptable only if we knew for certain that the same children were checked at both times.

Procedures to Reduce Bias, Enhance Consistency, and Check Accuracy

To reduce selection bias, we assigned two physician reviewers to screen and review titles/abstracts and full articles at every stage of the selection process. We assigned one physician and one health services researcher who was familiar with experimental design and biostatistics to abstract data. We assessed completeness of our collection of retrieved articles by cross-checking with studies included in other meta-analyses and references listed in review articles. The software program EndNote was used to check batches of articles added to the master list for duplicate references by comparing author, year title, and reference type. Following the importation of the first literature search, we used the software program EndNote to check subsequent references for duplication prior to their addition to the master list. After the master list was completed, we performed a second, manual, check to ensure no duplication.

To assess the extent of publication bias, we searched multiple sources and unpublished material identified by the Technical Expert Panel and internal content experts. We also studied

funnel plots—scatter plots of sample size versus the estimated effect size from each study. When publication bias existed, a portion of points would be missing from the funnel plot, typically at the null effect level. Because graphical evaluation can be subjective, we also conducted an adjusted rank correlation test (Begg, 1999) and a regression asymmetry test (Egger, Smith, Schneider et al., 1997) as formal statistical tests for publication bias. We conducted these tests using the statistical package Stata (StataCorp. 1999).

The mechanisms used to enhance consistency in screening and data abstraction include the use of pre-designed forms with explicit instructions and continuous and prompt resolution of discrepancies. Data were entered into a Microsoft Excel spreadsheet directly by the screeners or data abstractors. A third project staff cross-checked data for individual studies abstracted by each data collector. We resolved discrepancies by rechecking the article or by consensus via conference calls.

Preparation of Evidence Tables

An evidence table was prepared for each key question. Each evidence table provides a comprehensive tabular display of data abstracted from the literature in response to the question. It contains the name of the first author, year of publication, study design and quality score, how and by whom OME diagnosis was done, when and where the study took place, inclusion and exclusion criteria, important influencing factors, sample size, outcome measures and their definitions, and study findings. A total of four evidence tables was prepared; they are included in the section called Evidence Tables in this report.

Supplemental Analysis

Based on the discussions of technical experts during the conference calls and the designated time frame for the evidence assessment, a supplemental analysis plan was developed for each key question to synthesize the data.

Natural History

Question 1. What is the natural history (spontaneous resolution rate over time without treatment) for the following diagnostic groups: a) OME persisting after a discrete episode of acute otitis media, b) newly diagnosed OME of unknown duration (unilateral or bilateral), c) OME persisting for weeks or months (unilateral or bilateral), d) unilateral OME lasting 3 months or longer, and e) bilateral OME lasting 3 months or longer?

The outcome measures for this questions included complete and/or partial resolution rates, relapse/recurrence rates, and incidence of AOM after OME. The scope listed 31 non-treatment factors that might affect the course of the illness and confound the outcomes. They included: age, gender, ethnicity/race, socioeconomic status, number of hours in child care center, tobacco smoke exposure, season, number of children in household, breast-fed status, barotrauma challenges, OME laterality, hearing level, total duration of OME, age at first OM onset age of previous OME, number of previous OMEs, family history of OME, otitis prone (AOM), allergies, prior tubes, prior adenoidectomy, developmental delay, caregiver preference for treatment, caregiver education, examiner skill, examiner type, health care setting, monitoring time, monitoring frequency, monitoring personnel, and monitoring method.

Furthermore, the type of study is an important consideration for the assessment of natural history. A stratified random sample of a broad, well-defined population provides evidence of good generalizability, but may be restricted in the amount of clinical information on participants. A single (untreated) arm from a clinical trial will usually provide much more clinical evidence about OME, but this is usually assessed on a very selected group of children, making generalizing the results to the general population more difficult. For this evidence assessment we used only prospective cohort studies as these came closest to the ideal of enrolling a sample from a broad poulation.

The first step of the analysis was to obtain a distribution of studies stratified by the 5 diagnostic groups (namely, OME persisting after a discrete episode of acute otitis media, newly diagnosed OME of unknown duration (unilateral or bilateral), OME persisting for weeks or months (unilateral or bilateral), unilateral OME lasting 3 months or longer, bilateral OME lasting 3 months or longer), by type of outcome measures, and by non-treatment factors. This stratification provided us with an overview of the emphasis of past research in this area and an opportunity to identify gaps and areas for future research.

Using the DerSimonian and Laird random effects model (DerSimonian and Laird 1986) to pool rates across studies, we performed a meta-analysis on strata with more than 3 studies for a pooled estimate of an outcome with 95% confidence intervals. This method produces a summary measure by weighting each study's measure by the inverse of the sum of the within-study variance and the between-study variance. This approach allowed both sampling variation and between-study heterogeneity to affect the pooled estimate.

In addition to the pooled estimate, we reported the Q statistic and p-value for the Chi-squared test of heterogeneity, which tests the null hypothesis that the individual study results are homogeneous (Laird and Mosteller 1990).

Effects of Early-Life OM on Long-Term Speech and Language Development

Question 2. Do infants and preschool children with longer duration of early-life OME as compared to those with shorter duration OME have greater delays in speech and language development (receptive or expressive) later in life? One specific formulation of this question is: Is OME-associated conductive hearing loss in the first 3 years of life a risk factor for speech and language developmental delays? What are the risk factors that interact with the effect of OME on speech and language development in infants and preschool children?

For the first part of the question, the outcome of interest was speech and language developmental delay and the risk factor of interest was OME-associated conductive hearing loss and/or long versus short duration of early-life OME. For this question we included only comparative studies. Further, since prospective comparative cohort studies provide better evidence than retrospective comparative cohort studies, we conducted our assessment using only prospective comparative cohort studies.

The risk factor of interest was whether a child had or did not have OME-associated conductive hearing loss in the first 3 years of life, or whether duration of OME during the first three years of life was long or short. For this risk factor, we collected data on five related variables: hearing level, total duration of OME≥3 months, number of previous OMEs, duration of middle ear effusion (MEE), and repeated or persistent versus infrequent early-life OME. The hearing level was used to determine whether a child had OME-associated conductive hearing loss in the first 3 years of life. The total duration of OME greater or equal to 3 months was used

to define length of duration. We used the repeated or persistent versus infrequent early-life OME to define the risk. If a study did not classify the study subjects this way and (instead) reported data by number of previous OMEs and/or duration of MEE's, we sought the advice of the technical experts to stratify the samples based on these variables.

The influencing factors of outcome included both treatment and non-treatment factors. Here we are using "influencing factors" as a general term including risk factors for OM and/or confounding factors for the dependent variables of interest. The non-treatment factors included: age at first OM, gender, ethnicity/race, socioeconomic status, number of hours at a child care center, quality of child care, early intervention program, tobacco smoke exposure, number of children in household, breast-feeding status, OME laterality, allergies, developmental delay, OM complications, e.g. perforated TM, cholesteatoma, chronic illness of any type, caregiver education, quality of parent-child interaction, examiner skill, examiner type, health care setting, age at rechecks, frequency of rechecks, primary care provider, and type of equipment to measure hearing. Treatment factors included any combination of the following: tympanostomy tubes, adenoidectomy, myringotomy, antibiotics, systemic steroids, decongestant, antihistamine, N-acetyl-cysteine or others.

The outcome measures for this question related to speech and language developmental delay. These outcomes were measured by different instruments at different times, by different professionals, in different settings. In preparation for information synthesis, with the assistance of our speech and language technical expert, we classified the tests used in our final set of studies into the five outcome categories: expressive language, receptive language, expressive speech, receptive speech, and cognitive verbal intelligence. For analysis, we first stratified studies by the type of outcome measures, risk factor measures, and treatment or non-treatment risk factors. For any comparison among 3 or more studies, we conducted a meta-analysis. In each meta-analysis, we derived a pooled effect size defined as the proportion of standardized difference between the positive and negative otitis media groups. We pooled across studies the standardized mean differences between the groups and divided by a pooled standard deviation. We used a random effects model (DerSimonian and Laird, 1986) and the Hedges estimate of the pooled standard deviation (Hedges and Olkin, 1985). We used Stata (StatCorp. *Stata Statistical Software: Release 6.0.* College Station, TX: Stata Corporation. 1999) for the analyses.

To answer the second part of the question: "What are the risk factors that interact with the effect of OME on speech and language development in infants and preschool children?," we planned to conduct meta-regression analysis to identify the risk factors that contribute significantly to speech and language delays. For this analysis we would include both comparative and single cohort studies. Many technical issues must be addressed to set up data appropriately for meta-regression analysis. Due to restriction of the time frame, this part of the question was not included in this assessment but should be an area of future research.

Effects of Early-Life OM on Long-Term Hearing

Question 3. Do infants and preschool children with longer duration early-life OME as compared to those with shorter-duration OME have permanent (or sensorineural) hearing loss later in life? One specific formulation of this question is the following: Is OME-associated conductive hearing loss in the first 3 years of life a risk factor for permanent (or sensorineural) hearing loss later in life? What risk factors interact with the effect of OME on hearing loss later in life (unilateral or bilateral) in infants and preschool children?

The analysis plan for Question 3 followed that for Question 2. In preparation for information synthesis, we sought the assistance of our audiology technical expert, to group the tests used in our final set of studies into homogeneous categories.

Our audiology expert advised that (1) an acoustic reflex at 500 or 1000 Hz were both acceptable for study, (2) the criterion for abnormal reflex threshold would depend on the study, whether ipsilateral or contralateral, and on the frequency used for testing, (3) the abnormal reflex criteria should be based on normative data, and (4) quantitative tympanometry should be classified as:

a. Static Compensated Acoustic Admittance including: peak admittance, peak compensated admittance, peak compliance, static compliance, static admittance, and peak compliance.

b. Tympanometric Gradient including: gradient, pressure gradient and tympanometric gradient (Madsen compliance was excluded because Madsen compliance units were arbitrary units and the this instrument was from an era in which the units were not on a calibrated scale.)

c. Tympanometric Width referring to terms containing the words width, referring to tympanometry.

Diagnostic Methods for OME

Question 4. What are the sensitivity, specificity, and predictive values for alternative methods of diagnosing OME compared with one of the reference standards?

The diagnostic methods to be assessed included: a) signs/symptoms, b) non-pneumatic otoscopy, c) pneumatic otoscopy, validated or unvalidated examiner, d) binaural (or bilateral) micro-tympanoscopy, e) portable tympanometer, f) professional tympanometer, g) quantitative tympanometry, h) acoustic reflectometry, i) otoacoustic emissions, and j) audiometry, air or. bone conduction thresholds. The reference standards used to evaluate the accuracy of the diagnostic methods included: a) tympanocentesis, sedated or non-sedated, b) MRI, c) myringotomy, sedated or non-sedated, d) validated pneumatic otoscopy, and e) CT Scan.

Diagnostic methods based on algorithms, combinations of methods, or combination of scores, were not within the scope of this report because the sources of variation of such combinational methods would be difficult to detect in published articles and the analysis of them would not be feasible within our timeframe. Also excluded were studies where the experimental diagnostic test and the reference standard test were performed more than 24 hours apart.

Our strategy for evaluating the diagnostic value of a procedure was to derive pooled estimates for sensitivity, specificity, and prevalence rate for each diagnostic procedure and reference standard with 3 or more comparison studies. We used the DerSimonian and Laird random effects model (DerSimonian and Laird 1986) to derive random effects estimates and 95% confidence intervals. We also pooled the prevalence rates to determine the heterogeneity of the study populations. Using the pooled estimates, we plotted the performance of each diagnostic test in terms of sensitivity and (1-specificity) and identified the best performer among the tests included in the comparison. We then derived the positive and negative predictive values for the best diagnostic test for various prevalence levels.

To prepare for a meta-analysis for each comparison, we abstracted data from the evidence table; one meta-analysis was performed for sensitivity and specificity,. The following data

elements were entered into the SAS program to be converted into a SAS data set: study ID number, author and year of publication, number of adverse outcomes in the experimental group, total number of patients in the experimental group, number of adverse outcomes in the control group, and total number of patients in the control group. We used a SAS macro software program developed by RAND statistical staff to perform all meta-analyses and used the beta-test version of the software package "MetaGraphs" (1998, Belmont Research, Inc. 84 Sherman Street, Cambridge, MA 02140) for graphing.

The following statistics were generated from the SAS macro program: (a) study-level statistics (incidence rate, relative risk, risk difference, number needed to treat (NNT), odds ratio, and their 95 percent confidence intervals); (b) crude estimates and their 95 percent confidence intervals for all studies combined; (c) fixed effects estimates and their 95 percent confidence intervals for all studies combined; (d) random effects estimates and their 95 percent confidence intervals based on the DerSimonian and Laird method for pooling study results, and Chi-squared test of homogeneity; (e) weight for each study for both the fixed effects model and random effects model used to calculate of risk difference and relative risk.

To use MetaGraph for graphing, we entered the data into ASCII files using the UltraEdit-32 software. Funnel plots were produced for the purpose of screening possible publication bias, and the shrinkage plots were generated to display the effect size of each study and compare it against the overall model estimate, together with the 95 percent confidence limits. We evaluated the funnel plots graphically for asymmetry that resulted from the non-publication of small, negative studies. Because graphical evaluation can be subjective, we also conducted an adjusted rank correlation test (Begg, 1999) and a regression asymmetry test (Egger, Smith, Schneider et al., 1997) as formal statistical tests for publication bias. We conducted these tests in the statistical package Stata (StataCorp. 1999).

Identification of Peer Reviewers

At the beginning of the project, we requested nominations for technical experts and peer reviewers from 12 organizations. A total of 18 nominations were received for the Peer Review Panel. Experts in systematic reviews and meta-analysis were selected from a pool of experts associated with the Southern California Evidence-Based Practice Center but not involved with this project. The Project Staff, in consultation with the Task Order Officer, determined the relative mix of reviewers across the three domains (methodology, user, and clinical). In addition to domestic experts, we identified four European experts to serve as peer reviewers. The Peer Review Panel (**Table 15**) was composed of 18 members including family physicians, pediatricians, otolaryngologists, audiologists, speech-language pathologists, nurse practitioners, health planners, consumers, systematic review methodologists, statisticians, and non-U.S. experts in otitis media.

Peer Review Process

A copy of the draft evidence report was mailed to each peer reviewer on the panel, along with an instruction sheet (**Table 16**) for reviewing the draft evidence report. The Peer Review Panel was asked to respond within three weeks. Seventeen of the 18 peer reviewers responded with comments. A copy of the draft evidence report was also mailed to the members of the Technical Expert Panel and all technical experts responded with comments. Upon receipt of all responses from the peer reviewers and technical experts, the project staff compiled a summary of

the comments and changes and revised the draft evidence report. We forwarded all comments to the Task Order Officer for review. The peer reviewers' and technical experts' comments are included in **Appendix I,** together with the corresponding responses or actions taken by project staff.

Table 1: Technical Expert Panel

Technical Expert	Area of Expertise	Affiliation/Location
Larry Culpepper, MD, MPH	Family Medicine	Boston Medical Center, MA
Douglas G. Long, MD	Family Medicine	Manchester Community Health Center, NH
Richard M. Rosenfeld, MD, MPH	Otolaryngology	SUNY Health Science Center Brooklyn, NY
Norman Wendell Todd, Jr., MD	Otolaryngology	Emory University, GA
Allan Lieberthal, MD	Pediatrics	Southern California Kaiser Permanente Medical Group, CA
Anthony Magit, MD	Pediatric Otolaryngology	Children's Hospital San Diego, CA
Jack Paradise, MD	Pediatrics	Children's Hospital, Pittsburgh, PA
Ross Miller, MD	Quality Management	CIGNA Health Care, CA
Joanne Roberts, PhD	Speech and Hearing	University of North Carolina, NC
Lisa L. Hunter, PhD	Audiology	University of Minnesota, MN
Linda Carlson, MS, RN, CPNP	Nurse Practitioner	Statesboro, GA
Fran Goldfarb, MA	Consumer	Family Voices, Los Angeles, CA

Table 2: Questions Suggested for Consideration in Evidence Report

1. What is the relative risk of developing OME in the child who has food or inhalant allergies compared to the child without food or inhalant allergies

2. What is the natural history for various types of OME?

3. What is the long-term level of speech and language development in infants and preschool children with untreated OME? What are the high risk groups?

4. What is the accuracy of various diagnostic methods?

5. When should conservative treatment (non-surgical) be considered a failure?

6. What is the evidence on effectiveness of various diagnostic instruments in deciding on intervention for OME?

7. What is the evidence regarding level of hearing decrease and whether unilateral or bilateral hearing decrease is an indication for intervention?

8. What is the effectiveness of the use of hearing levels to decide on intervention for OME?

9. Are antibiotics more effective than placebo in treating OME?

10. Are steroids more effective than placebo in treating OME?

11. Do antibiotics add an incremental benefit to steroids in treating OME?

12. Are interventions for allergies (food or inhalant) more effective than placebo in treating OME?

13. Are antihistamines and/or decongestants more effective than placebo in treating OME?

14. Are tympanostomy tubes more effective than other interventions in treating OME?

15. Is adenoidectomy more effective than other interventions in treating OME of greater than 3 months duration?

16. Is tonsillectomy more effective than other interventions in treating OME of greater than 3 months duration?

17. Is myringotomy more effective than other interventions in treating OME of greater than 3 months duration?

18. Are alternative or complementary therapies more effective than other interventions in treating OME of greater than 3 months duration?

19. Are prophylactic antibiotics more effective than other interventions in treating OME?

20. What is the effectiveness of monitoring by pneumatic otoscopy, tympanometry, acoustic reflectometry with spectral gradient, and otoacoustic emissions to decide on intervention for OME?

Table 3: Ranking of Potential Key Questions

Of the 20 questions, each technical expert ranked the top 10 questions **from 10 (highest priority) to 1 (lowest priority),** (experts' identification numbers were randomly assigned).

	Topic of question	**Rank Total**	E 1	E 2	E 3	E 4	E 5	E 6	E 7	E 8	E 9	E 10	E 11	E 12
1.	Food or inhalant allergies	**8**				5					3			
2.	Natural history	**78**	10	9	10	9	7	9	4	2		6	7	5
3.	Speech and language development	**97**	8	10	4	8	9	5	8	5	10	10	10	10
4.	Accuracy of diagnostic methods	**57**			5	10	6	10	6	9		8	1	2
5.	Conservative treatment	**39**	9	5	6	4	1		3			1	9	1
6.	Diagnostic instruments vs intervention	**38**	7	8				6		9		8		
7.	Level of hearing decrease	**67.5**		7		7	8	4	9	6.5	9		8	9
8.	Hearing levels and intervention	**16.5**		6	1			3		6.5				
9.	Antibiotics versus placebo	**51**	6	4	9	6	5		1		7	5	2	6
10.	Steroids versus placebo	**23**					4		7	1	6		5	
11.	Antibiotics and steroids	**22**	5	2	3		3				5		4	
12.	Treatment for allergies vs placebo	**2**	1			1								
13.	Antihistamines/decongestants vs placebo	**10**							10					
14.	Tympanostomy tubes vs other interventions	**51**	3	1	8		10	2	2	4	4	3	6	8
15.	Adenoidectomy vs other interventions	**32**			7		2	1	5	3	2	2	3	7
16.	Tonsillectomy vs other interventions	**2**	2											
17.	Myringotomy vs other interventions	**5**			2	3								
18.	Alternative/complementary therapies vs other interventions	**10**				2		7			1			
19.	Prophylactic antibiotics vs other interventions	**15**	4	3								4		4
20.	Effectiveness of diagnostic methods for monitoring	**36**						8		9	8	8		3

Note: Kendall Coefficient of Concordance = 0.36, p=0.0001.

Table 4: Comments from Technical Expert Panel on Potential Key Questions

Potential Key Questions		Comments/Notes
1.	What is the relative risk of developing OME in the child who has food or inhalant allergies compared to the child without food or inhalant allergies?	• Difficult to obtain evidence • Doesn't seem to have adequate research base • Assumed to be a minor player, but there could be a surprise here
2.	What is the natural history for various types of OME?	
3.	What is the long-term level of speech and language development in infants and preschool children with untreated OME? What are the high risk groups?	• Evidence incomplete • Several new prospective studies are available • This is probably the most important single question for justifying surgical intervention • Available results inconsistent and/or contradictory. No definitive answers yet available
4.	What are the accuracy of various diagnostic methods?	• Combine with 6 and 20 • Probably not much data there • Add binocular micro-tympanoscopy, MRI, and quantitative tympanometry to the list of methods of diagnosing OME. Binocular micro-tympanoscopy, which I do to every patient that I assess in the clinic, is using the surgical "operating" microscope to view each tympanic membrane. I think MRI should be the "gold standard" (see Swarts JD et al.: In vivo observation with magnetic resonance imaging of middle ear effusion in response to experimental underpressures. Ann Otol Rhinol Laryngol 104:522-528, 1995). To diagnose OME by looking at the tympanic membrane is traditional, but provides only a look at the "window on the middle ear". To use fluid found at myringotomy as the "gold standard" assumes that either (1) no inhalational general anesthetic has been given to the patient, or (2) there is no effect of inhalational general anesthetic on the finding of fluid in that part of the mesotympanum when the myringotomy is done. Since practically, inhalational general anesthetic is usually administered to patients getting myringotomy in the US in 1999, I'm not comfortable with this first assumption. From clinical experience, I've very uncomfortable with the second assumption. MRI also affords an opportunity to measure the volume of the middle ear system (I'm including mastoid, mesotympanum, epitympanum and all the spaces normally containing "gas" in this definition of the middle ear system); smallness of volume of the middle ear system is a well-documented correlate of severity of otitis media condition, and also a correlate of the severity of the anatomic eustachian and skull base differences of otitis media.) [The MRI data would also help, I think, in answering "Potential Key Question # 2.] Quantitative tympanometry is a promising technique. ANSI mandated that all new tympanometers (as of 1996, as I recall) provide quantitative information. Such quantitative information, [DeChicchis & Todd, unpublished data] to date show advantages over the qualitative A-B-C Jerger classification.

Table 4 (Continued)

5.	When should conservative treatment (non-surgical) be considered a failure?	• Unclear question • Very broad category • Unclearly written. What is the conservative treatment • No definitive information on which to base an answer
6.	What is the evidence on effectiveness of various diagnostic instruments in deciding on intervention for OME?	• Same as 4 • Combine with #4 and #20
7.	What is the evidence regarding level of hearing decrease and whether unilateral or bilateral hearing decrease is an indication for intervention?	• Minimal evidence • Combine with 8 • Pairs with question 3 • No definitive information on which to base an answer
8.	What is the effectiveness of the use of hearing levels to decide on intervention for OME?	• Same as 7 • Combine with 7
9.	Are antibiotics more effective than placebo in treating OME?	• Is this still an issue? • Is this still an issue? Has it been addressed with meta-analysis already? • Let's see if there is any new data out there
10.	Are steroids more effective than placebo in treating OME?	• There wasn't enough data during the original panel. Is there now?
11.	Do antibiotics add an incremental benefit to steroids in treating OME?	• This area showed some potential promise last time • Should be "Do steroids add an incremental benefit to antibiotics"
12.	Are interventions for allergies (food or inhalant) more effective than placebo in treating OME?	• Minimal evidence • Inadequate research base
13.	Are antihistamines and/or decongestants more effective than placebo in treating OME?	• If used need to indicate if there is a presence or absence of allergies • Still an issue? Addressed in last OME guidelines
14.	Are tympanostomy tubes more effective than other interventions in treating OME?	• Good to test the current recommendation to see if they still have validity • Answers might vary depending on outcome measures
15.	Is adenoidectomy more effective than other interventions in treating OME of greater-than 3 months duration?	• Need to indicate in what age group and if used combine with myringotomy • The key is age of patient. Are there studies for younger children? • Answers might vary depending on outcome measures • Though I think adenoidectomy has some benefit to some patients in helping the resolution of their otitis media, none of the purported mechanisms make any sense to me. One purported mechanism is to decrease the "cesspool" of the nasopharynx; but,

Table 4 (Continued)

		cystic fibrosis patients, who typically have the worse cesspools of any nasopharynges, have (on average) less otitis problems than the general population. Another purported mechanism is removing the mass of lymphoid tissue in the nasopharynx; but, I don't know of any data that size of adenoid tissue correlates with either the occurrence or severity of otitis media. Indeed, in my experience with a population who have one of the highest rates of otitis of any population, adenoid tissue is usually scant. I'd love to know the explanation by which adenoidectomy improves the course of otitis media in some children. (I suspect the explanation is scarring in the adenoid bed, that stabilizes the posterior lamina of the eustachian cartilage. If this is indeed the explanation, then it is one of the rare occasions that surgery benefits a patient by inducing scarring.)
16.	Is tonsillectomy more effective than other interventions in treating OME of greater-than 3 months duration?	• Still an issue?
17.	Is myringotomy more effective than other interventions in treating OME of greater-than 3 months duration?	• Still an issue
18.	Are alternative or complementary therapies more effective than other interventions in treating OME of greater-than 3 months duration?	• Minimal evidence • Inadequate research base? • Probably won't find much, but the public will be clamoring for it
19.	Are prophylactic antibiotics more effective than other interventions in treating OME?	• Still an issue
20.	What is the effectiveness of monitoring by pneumatic otoscopy, tympanometry, acoustic reflectometry with spectral gradient, and otoacoustic emissions to decide on intervention for OME?	• Similar to 6 combined • Combine with #6 and #4 • Add binocular micro-tympanoscopy, MRI, and quantitative tympanometry to the list of methods of diagnosing OME. See my comments about "Potential Key Question" 4.
	Epidemiology	• The material on epidemiology thus far circulated omits mention of a key factor in predisposing infants and children to otitis media, namely, low socioeconomic status. For data and a discussion, see our report in Pediatrics 1997;99:318-333.
	Definition of AOM	• I disagree with the definition of AOM as stated in your recent Definition section. As stated, the definition would call for a diagnosis of AOM in a child with middle-ear effusion and rapid onset of either irritability or fever. On the one hand, some infants with AOM have none of the 4 listed signs or symptoms, and the diagnosis is made on the basis of specific tympanic membrane findings--bulging and/or marked erythema--in addition to findings of middle-ear effusion. On the other hand, some infants with only OME rather

Table 4 (Continued)

		than AOM present with rapid onset of fever and/or irritability that may be due the underlying viral respiratory tract infection. If such patients are assumed to have AOM, much unnecessary antibiotic will be prescribed. The point to be made is that diagnosis should not be linked to either the presence or the absence of signs or symptoms such as fever and irritability that are nonspecific. Ear pain, on the other hand, is reasonably specific. This issue was discussed in Commentary in Pediatrics 1995;96:712-715.

Table 5: Overall Causal Pathway for OME

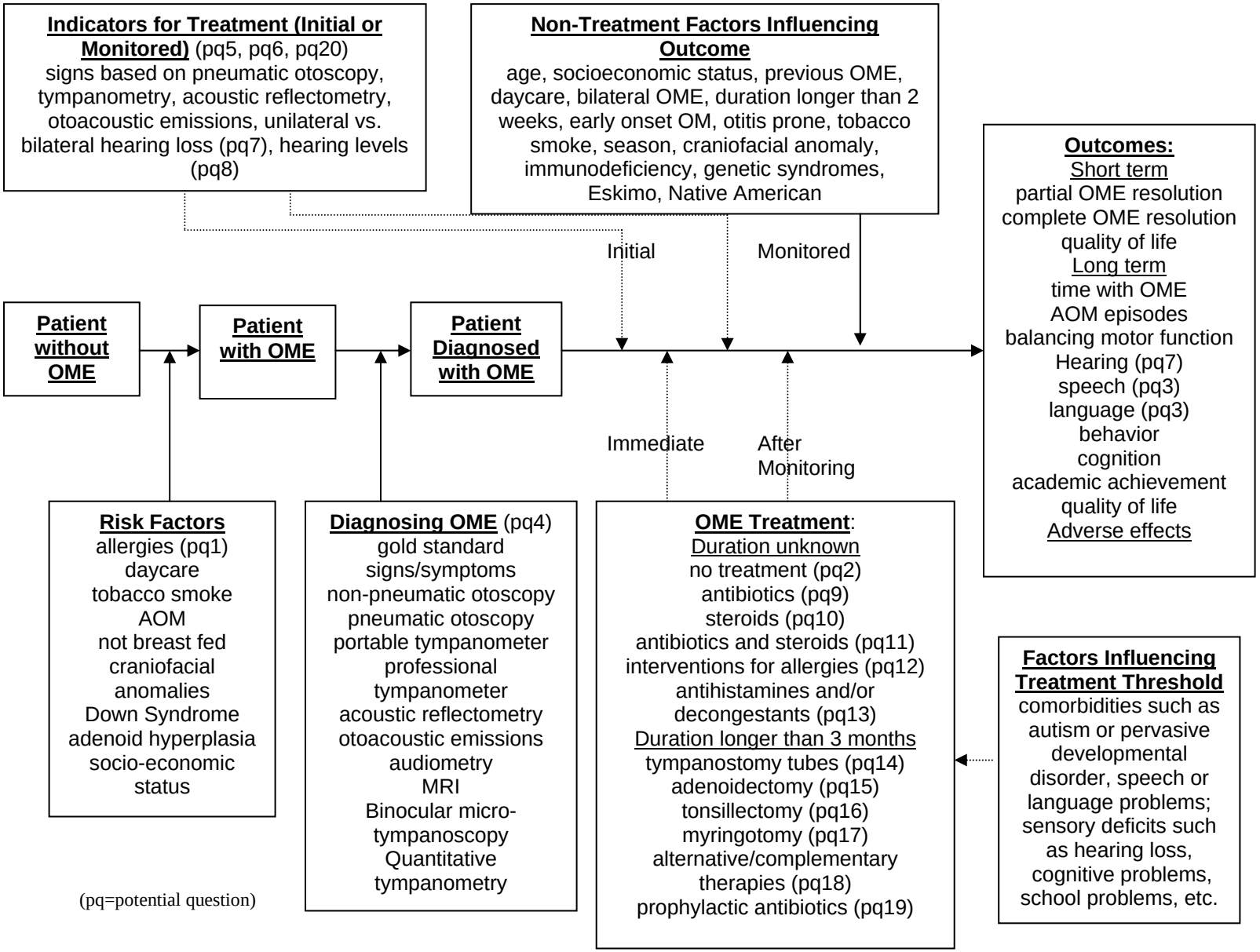


Table 6: Causal Pathway for Key Question 1 on Natural History

What is the natural history (spontaneous resolution rate over time without treatment) for:
- OME persisting after a discrete episode of acute otitis media,
- newly diagnosed OME of unknown duration (unilateral or bilateral),
- OME persisting for weeks or months (unilateral or bilateral),
- unilateral OME lasting 3 months or longer,
- bilateral OME lasting 3 months or longer?

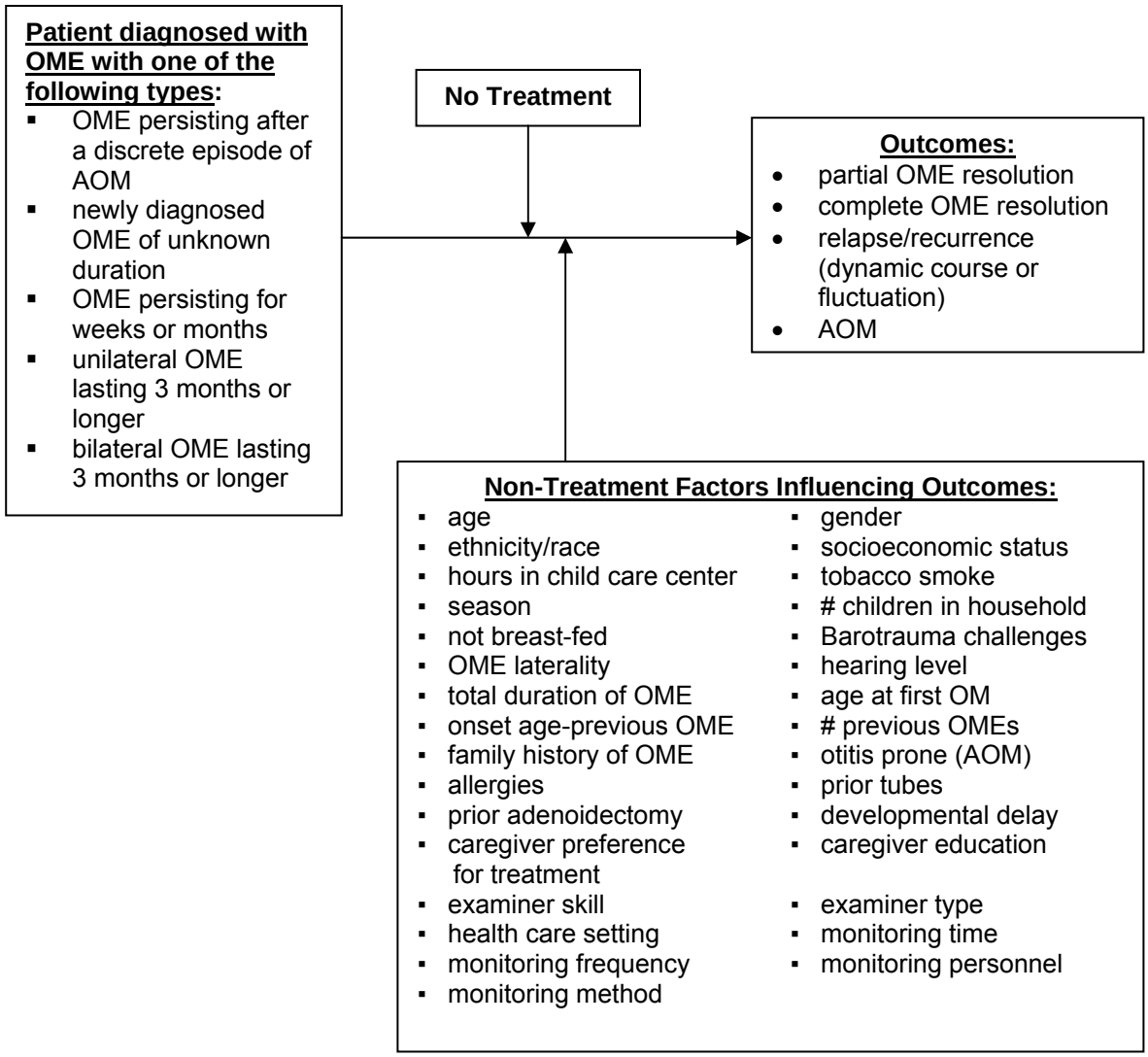


Table 7: Causal Pathway for Key Question 2 on Speech and Language

a) Do infants and preschool children with longer duration early life OME as compared to those with shorter duration OME have greater delays in speech and language development (receptive or expressive) later in life? One specific formulation of this question is: Is OME-associated conductive hearing loss in the first 3 years of life a risk factor for speech and language developmental delays?

b) What are the risk factors that interact with the effect of OME on speech and language development in infants and preschool children?

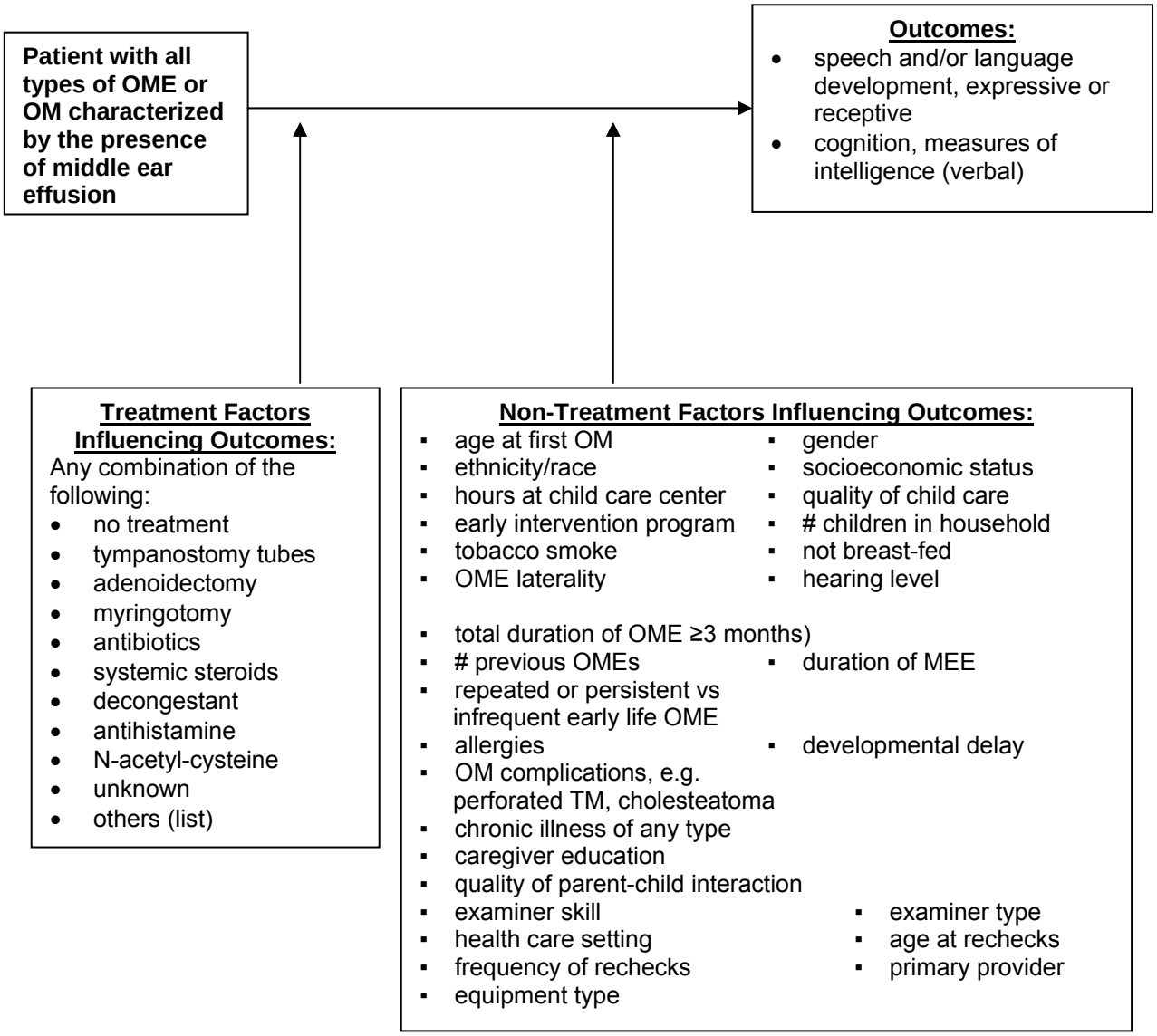


Table 8: Causal Pathway for Key Question 3 on Hearing

a) Do infants and preschool children with longer early-life OME as compared to those with shorter duration OME have permanent (or sensorineural) hearing loss later in life? One specific formulation of this question is "Is OME-associated conductive hearing loss in the first 3 years of life a risk factor for permanent (or sensorineural) hearing loss later in life?"

b) What are the risk factors that interact with the effect of OME on hearing later in life (unilateral or bilateral) in infants and preschool children?

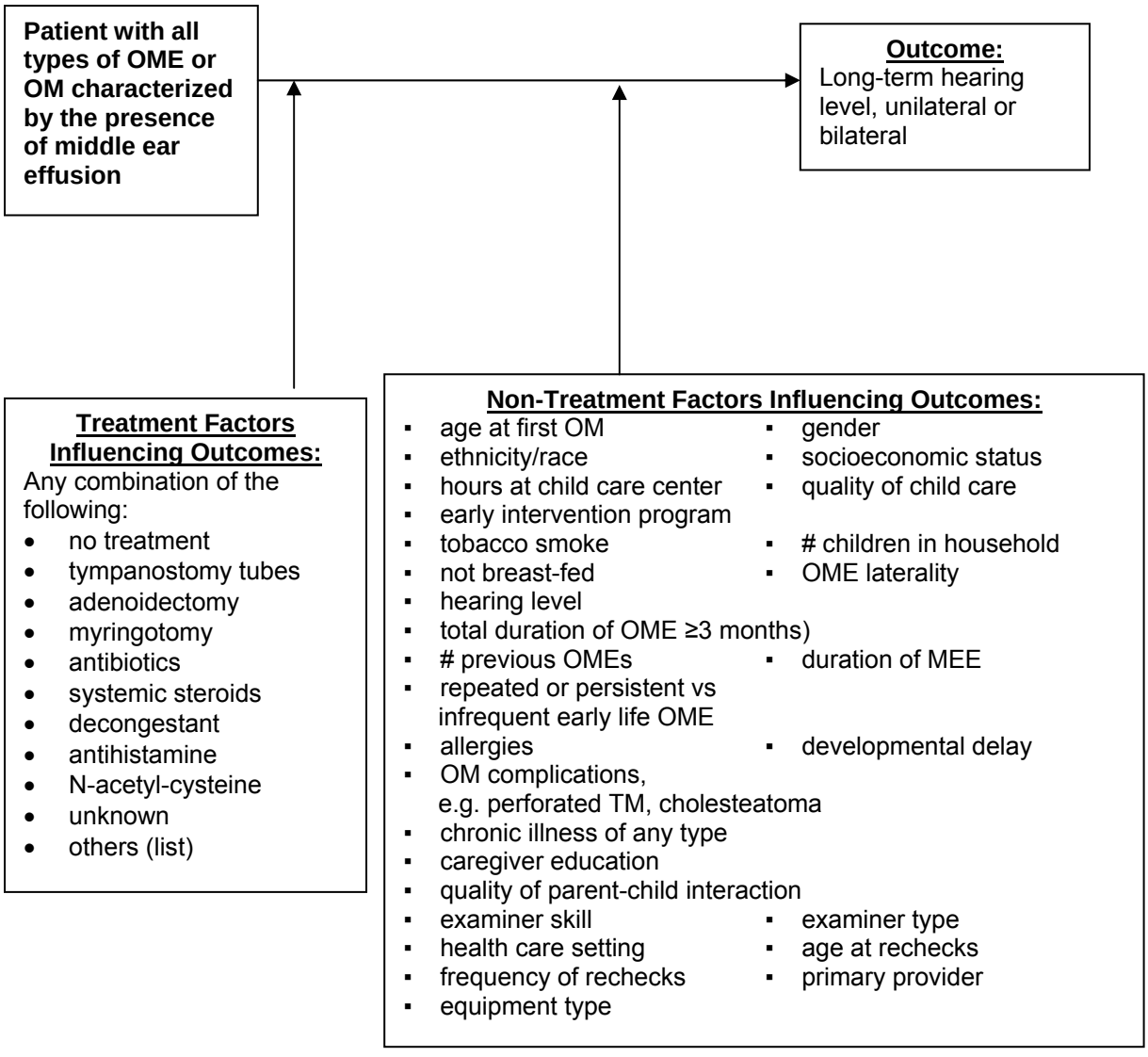


Table 9: Causal Pathway for Key Question 4 on Diagnostic Methods

What are the sensitivity, specificity, and predictive values for alternative methods of diagnosing OME compared with one of the gold standards?

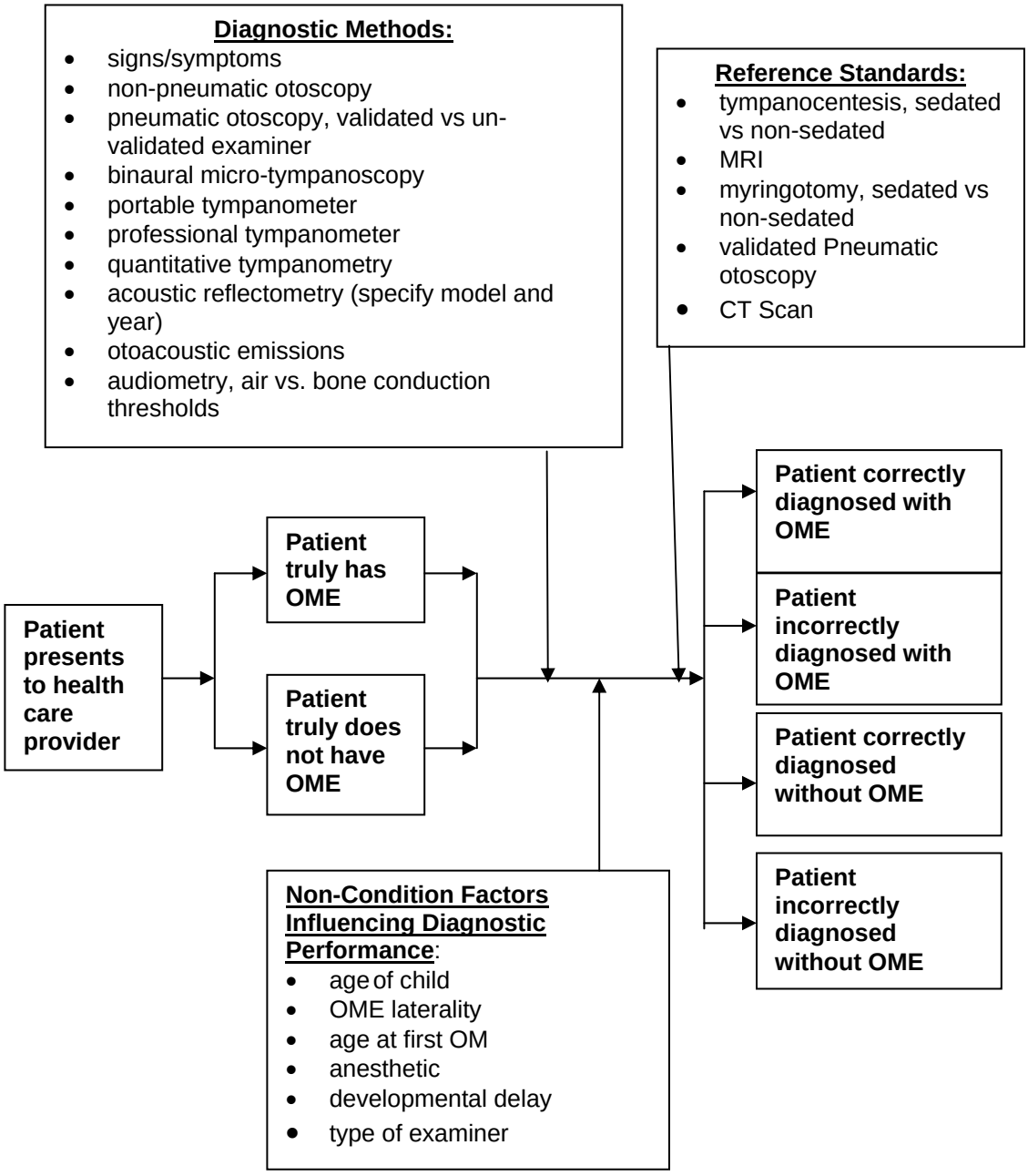


Table 10: Ranking of Influencing Factors by Technical Experts

Key Question 1: Natural History

Factors	Total "Yes" Based on Responses of 11 Technical Experts[a]
total duration of OME (≥3 mos)	10
otitis prone (AOM)	9
number of previous OMEs	9
number of hours attending child care center	9
tobacco smoke exposure	9
season of the year	9
age at first OM	9
not breast-fed	9
allergies	8
age of child	8
family history of OME	7
number of children in household	7
prior tubes	7
ethnicity/race	6
barotrauma challenges	6
prior adenoidectomy	5
socioeconomic status	5
laterality, unilateral versus bilateral	4
gender	3
skill to diagnose (validated)	3
age of onset of previous OME	3
tympanometry	3
monitoring frequency	2
monitoring time	2
MRI	2
type of examiner	2
setting of care	1
parent/caregiver education	1
hearing level, conductive versus sensorineural	1
primary provider	1
acoustic reflectometry	1
otoscopy	1
pneumatic otoscopy	1
parent/caregiver preference for treatment	1
developmental delay	0

[a] 11 technical experts responded; 1 abstained.

Table 10 (Continued)

Key Questions 2: Speech and Language

Factors	Total "Yes" Based on Responses of 12 Technical Experts
developmental delay	11
quality of child care	10
hearing level, conductive versus sensorineural	10
parent/caregiver education	10
quality of parent-child interaction	10
socioeconomic status	8
laterality, unilateral versus bilateral	8
early intervention program	7
total duration of OME (≥3 mos)	7
gender	6
number of children in household	6
duration of middle ear effusion	6
chronic illness of any type	6
number of hours attending child care center	5
number of previous OMEs	4
presence of active ear disease	4
ethnicity/race	3
tobacco smoke exposure	3
OM complications	3
child temperament	3
allergies	2
ambient noise	2
age at first OM	1
not breast-fed	1
skill to diagnose (validated)	1
type of examiner	0
setting of care	0
recheck times	0
frequency of recheck	0
primary provider	0
tympanometry	0
acoustic reflectometry	0
pneumatic otoscopy	0
MRI	0
equipment	0
audiometry	0

Note: Items in bold were added in the second poll after the first poll.

Table 10 (Continued)

Key Questions 3: Hearing

Factors	Total "Yes" Based on Responses of 12 Technical Experts
OM complications	10
laterality, unilateral versus bilateral	7
hearing level, conductive versus sensorineural	7
developmental delay	6
presence of active ear disease	6
total duration of OME (≥3 mos)	4
number of previous OMEs	3
duration of middle ear effusion	3
chronic illness of any type	3
ambient noise	3
allergies	2
age at first OM	1
ethnicity/race	1
socioeconomic status	1
quality of child care	1
early intervention program	1
tobacco smoke exposure	1
number of children in household	1
child temperament	1
equipment	1
audiometry	1
gender	0
number of hours attending child care center	0
not breast-fed	0
parent/caregiver education	0
quality of parent-child interaction	0
skill to diagnose (validated)	0
type of examiner	0
setting of care	0
recheck times	0
frequency of recheck	0
primary provider	0
Tympanometry	0
acoustic reflectometry	0
pneumatic otoscopy	0
MRI	0

Note: Items in bold were added in the second poll after the first poll.

Table 10 (Continued)

Key Question 4: Diagnostic Tests

Factors	Total "Yes" Based on Responses of 11 Technical Experts[a]
age of child	11
otolaryngologist	6
nurse practitioner	5
pediatrician	5
family physician	5
laterality, unilateral versus bilateral	5
anesthetic	5
age at first OM	5
developmental delay	4
physician assistant	4
others	3

[a] 11 technical experts responded; 1 abstained.

Table 11: Reason for Rejection of Titles/Abstracts at Initial Screening (N=3,200)

Reason Code	Reason for Rejection	Number (Percent) of Citations
R0	Written in non-English language	170 (5.3%)
R1	Case report/editorial/letter/ clinical/overview/practice guidelines/consensus statements	459 (14.3%)
R2	Non-human subjects	19 (0.6%)
R3	Study condition not OM	804 (25.1%)
R4	Age of study population >12 years[a]	57 (1.8%)
R5	Study population exclusively on any one of the following: Craniofacial defects, primary mucosal disorders, Immunodeficiencies, or Down or other genetic syndromes	15 (0.5%)
R7	Any key questions not addressed	697 (21.8%)
R8	Duplicate citation	9 (0.3%)

[a] The age limit was later extended to 22 years of age for Questions 2 and 3.

Table 12: Reason for Rejection of Full-Length Articles at Secondary Review

Reason Code	Reason for Rejection	Cochrane and Medline N=614	AAP Data Files N=29	Symposia on OME N=107	Technical Experts N=17	Articles and books N=31	Total N=798[a]
R0	Non-English language	1	0	0	0	0	1 (0.1%)
R1	Case report/editorial/ letter/clinical/ overview/ practice guidelines/ consensus statements	142	24	22	4	17	209 (26.2%)
R2	Non-human subjects	4	0	0	0	2	6 (0.8%)
R3	Study condition not OM	33	0	0	2	2	37 (4.6%)
R4	Age of study population >22 years	0	0	2	0	1	3 (0.4%)
R5	Study population exclusively one of the following: craniofacial defects, primary mucosal disorders, immunodeficiencies, or Down or other genetic syndromes	3	0	1	2	0	6 (0.8%)
R7	Key questions not addressed	354	5	74	9	2	444 (55.6%)
R8	Duplicate citation	39	0	3	0	7	49 (6.1%)
R9	Data not abstractable from article	38	0	5	0	0	43 (5.4%)

[a] 3 incorrect citations not included.

Table 13: Results of Secondary Screening of Full-Length Articles

	Cochrane and Medline	AAP Data Files	Symposia Proceedings on OME	Technical Experts	Articles and books	All Sources
Total Citations	3200	1441	159[a]	39	40	4879
Number of articles reviewed at secondary screening	975[b]	32	159	36	40	1242
Number accepted after secondary screening	372	3	52	19	5	451
Question 1	127	0	14	0	0	141
Question 2	74	1	20	17	0	112
Question 3	157	2	22	2	3	186
Question 4	69	0	4	0	2	75
Total of above [c]	427	3	60	19	5	514

[a] Exact number of citations not determined.
[b] Four cases previously rejected because of age limit >12 were added to the original 971 accepted citations.
[c] The 'total of above' number can exceed the number accepted because an article can address more than one question.

Table 14: Results of Tertiary Screening of Full-Length Articles for Analysis

	Accepted at Secondary Screening	Abstracted in Evidence Tables
	449	114 (25%)
Question 1	141	38 (27%)
Question 2	112	21 (19%)
Question 3	186	8 (4%)
Question 4	75	52 (69%)
Total of above [a]	514	119 (23%)

[a] An article can address more than one question.

Table 15: Peer Review Panel

Peer Reviewer	Area of Expertise	Affiliation/Location
Howard Bauchner, M.D.	General Pediatrics	Child and Adolescent Health Scholar in Residence Agency for Healthcare Research and Quality
Hanan S. Bell, Ph.D.	Methodology reviewer	Seattle, WA
Alfred O. Berg, MD, MPH	Family Medicine	University of Washington, Seattle, WA
Patricia A. Fall, MS, CRNP	Nurse Practitioner	Wexford, PA
George A. Gates, MD	Otolaryngology	University of Washington, Seattle, WA
Janice Goertz, RN, CPNP	Nurse Practitioner	Portage, MI
Judith Gravel, PhD	Hearing and Speech	Albert Einstein College of Medicine, Bronx, NY
Mark P. Haggard, Ph.D.	Hearing/ Psychoacoustics	Institute for Hearing Research, Nottingham, UK
Vic Hasselblad, Ph.D.	Meta-analysis reviewer	Duke University, Durham, NC
Tracy Lieu, MD	Pediatrician/Health Plan	Harvard Pilgrim Health Care, Boston, MA
Martin C. Mahoney, MD, Ph.D.	Family Medicine	DeGraff Family Medicine, North Tonawanda, NY
A. Richard Maw MS FRCS	Otolaryngology	Bristol Royal Infirmary, Bristol UK
Robert Ruben, MD	Otolaryngology	Montefiore Medical Center, Bronx, NY
Anne GM Schilder, MD, Ph.D.	Otolaryngology	University Medical Center Utrecht, The Netherlands
Steve Shelov, MD	Pediatrics	Scarsdale, N.Y
Sylvan Stool, MD	Otolaryngology	The Children's Hospital, Denver, CO
Robin Yurk, MD, MPH	Consumer/Health Plan	Community Clinic, Inc. Rockville, MD
Dr. J.O.M. Zaat (Joost Zaat)	Methodology reviewer	Purmerend, The Netherlands

Table 16: Instructions for Reviewing Draft Evidence Report

Enclosed is a draft evidence report on the diagnosis and treatment of otitis media with effusion. You may make your comments either directly on the draft evidence report, or on a separate sheet of paper. If you choose to record your comments on a separate piece of paper, please use the page and paragraph number to identify to which part of the report your comments pertain.

We ask that you consider the following questions while you read this report. We realized that some of the questions may not pertain to your area of expertise. Please feel free to comment only on those that you feel most suited to answer.

1. **Overall evaluation**

Is it clear what we did? You may agree or disagree with our methods, findings, or conclusions, but you should be able to understand what we did in order to produce this report.

2. **Methodology**

Are the methods we used appropriate:
a) for identifying the key questions of interest from the panel of technical experts?
b) for searching and reviewing the identified literature?
c) for synthesizing the literature?

3. **Evidence**

a) Did we miss any crucial pieces of information in our literature search?
b) Does the evidence support the conclusions?

4. **Utility**

Would you find this information to be useful if you had to develop clinical practice guidelines or medical review criteria for diagnosis and treatment of otitis media with effusion in children?

Chapter 3. Results

Key question 1: What is the Natural History (Spontaneous Resolution Rate over Time without Treatment) for

- OME persisting after a discrete episode of acute otitis media,
- newly diagnosed OME of unknown duration (unilateral or bilateral),
- OME persisting for weeks or months (unilateral or bilateral),
- unilateral OME lasting 3 months or longer,
- bilateral OME lasting 3 months or longer?

Literature Review

After initial screening of 4,879 titles or abstracts, we identified 449 articles for review. After secondary screening of the 449 articles, we identified 141 articles that fell within the scope of this question. Tertiary screening identified 38 articles on prospective cohort studies for potential abstraction. After reassessing articles included in the systematic review by Rosenfeld (1999), we included three more articles for potential abstraction. We eliminated one article on r OME following acute otitis media which was addressed in a recent evidence analysis. A total of 26 prospective cohort studies and one retrospective-prospective cohort study were identified among these 40 articles. **Table 17** lists the studies and cohorts examined. Abstraction was possible from 33 of the 40 articles. **Table 18** lists the studies referred to by Rosenfeld (1999) and their disposition in the present evidence-based analysis. **Table 19** lists the articles excluded because relevant data could not be abstracted.

Findings

Evidence Table 1 presents the study characteristics, population characteristics, risk factors, and findings for the 27 cohort studies described in 34 articles that responded to this question**.** **Table 20** presents the study quality scores for the studies included in **Evidence Table 1**. The quality scores (see Methods) for these studies fell at the low end of the possible range of 1(lowest) –to 6(highest). Three of the cohort studies had a score of 4, sixteen had a score of 3, seven had a score of 2, and one had a score of 1.

Although we accepted these 27 studies for the natural history analysis, half of them (thirteen) failed to document that the subjects had received no medical or surgical treatment during the course of the study that could affect the outcome of OME. In the three studies that claimed that their subjects received no treatment, the investigators failed to document how adherence was maintained or confirmed. In studies of eleven cohorts, the authors mentioned that children received antibiotics or underwent surgical procedures that could affect OME outcome. In those studies that reported numbers of children who received treatment, the proportion was relatively

small compared to the total number studied. Except in two studies, results were not stratified by treatment condition, even in those articles that reported treatments received by study subjects. Further, the majority of studies, but not all, used tympanometry as the sole diagnostic test of OME. However, the criteria for each tympanogram type in these studies are similar (**Table 21**). Table 21 lists the pressure and immittance parameters for tympanogram types for those studies that used tympanometry as the sole diagnostic criteria for OME.

Tables 22 and 23 provide OME resolution rates by ear, as reported in the majority of the studies, and by child for those studies in which data could be stratified by unit of analysis (ear vs. child), age group, and OME type respectively. For clarity, we have listed only those resolution intervals with onset at the study's inception. For clarity, when tympanometry was the diagnostic method, counts for tympanogram type B transition to A are shown when provided, while other tympanogram transitions are shown in **Evidence Table 1.**

Data from a number of studies were reported in a format that made them unusable for quantitative syntheses. Ten studies were eliminated from quantitative syntheses, because they failed to stratify by age (i.e. less than and greater than 3 years old). Day-to-day variability in tympanogram types was described in a cohort of kindergarten children examined on each weekday for 30 days (Moller and Tos,1990). Ernston and Sundberg (1984) described a group of children who participated in a controlled trial of children with OME that persisted for at least 3 months; they found that 15.3 percent (11 of 72) of such children had OME resolution by five weeks followup. A study of children with OME that persisted for 3 months or more, showed 45 percent (49 of 109) OME resolution at 2.5 year followup (Leiberman and Bartal,1986). However, the investigators acknowledged that middle-ear effusion noted after such a long interval could be either persistent or recurrent. Birch and Elbrønd (1984) and Zielhuis, Rach, and van den Broek (1990) derived equations to describe the OME resolution rates they observed in their cohorts, but we were unable to abstract actual counts from these articles.

We performed two sets of meta-analyses that matched age groups, unit of analysis, outcome type, and time to resolution in three or more cohorts. All meta-analyses presented here used the 'ear' as the unit of analysis. Few studies considered the child or the episode as the unit of analysis and no meta-analyses were possible.

Resolution at 6-Week Followup The first set of meta-analyis contains two meta-analyses (**Table 24**). The meta-analyses showed that if the criteria for resolution were tympanogram type B or C transition to A or by otoscopy, 42.3 percent (95% CI: 24.1%, 60.6%) of ears with newly diagnosed OME of unknown duration in children older than 3 years had resolution by the 6-week exam. If the criteria for resolution were modified to tympanogram type B transition to A or otoscopy, the proportion of ears in children older than 3 years old with resolution at 6-week followup was 37.2 percent (95% CI: 1.8%, 72.5%). Spontaneous resolution rates were significantly different among the cohorts for both definitions. In these studies, the OME resolution rates were calculated by determining the proportion of children without OME at followup, whether or not their baseline OME had resolved and recurred at an earlier point. Thus, these are not cumulative resolution rates (Sly, Zambie, Fernandes et al., 1980; Lamothe, Boudreault, Blanchette et al., 1981).

Resolution at 3-Month Followup The second set of meta-analyses are shown in **Table 25**. The first two meta-analyses assessed resolution of OME by ear by the time of 3-month followup in children older than 3 years. If the criteria for resolution were tympanogram type B or C transition to A, the proportion of ears with resolution of newly diagnosed OME of unknown duration was 42.7 percent (95% CI: 29.3%, 56.1%) among children older than 3 years. If the

criteria for OME resolution were tympanogram type B transition to A, then the proportion of ears with resolution was 22.5 percent (95% CI: 5.9%, 39.0%). Spontaneous resolution rates were not significantly different among the studies in either comparison, except for Tos, Holm-Jensen, Sörenson, and Mogensen (1982) which had lower resolution rates. Because Fiellau-Nikolajsen and Lous (1979), Fiellau-Nikolajsen (1979), and Renvall, Anniansson, and Lidèn (1982) terminated followup whenever a child had a type A or normal tympanogram and because Tos, Holm-Jensen, Sörenson, and Mogensen (1982) provided data for calculation of the cumulative OME resolution rate, this estimate represents a cumulative resolution rate. The last three meta-analyses in Table 25 show the derivation of the cumulative resolution rates.

Resolution Rates for Younger Age Groups Only two studies examined the resolution rates for each of the age groups of less than 6 months and 3-months to 3-years (see **Tables 22 and 23**). Thus, no meta-analyses for children under 3 years of age were performed.

The Role of Influencing Factors in Resolution Similarly, because very few studies assessed the role of factors that might influence resolution, no meta-analyses were performed. The results of individual studies are summarized here:

- Lamothe, Boudreault, Blanchette, and colleagues (1981) and Fiellau-Nikolajsen and Lous (1979) assessed the effects of gender and found quicker resolution among females than males.

- Fiellau-Nikolajsen (1979) noted that children who received at-home care had quicker resolution of OME than did children in daycare.

- Sly, Zambie, Fernandes, and colleagues (1980) compared OME resolution in small cohorts first studied in February with those first studied in September and found resolution to be more rapid in the February cohort.

- Lamothe, Boudreault, Blanchette, and colleagues (1981) also assessed the effect of the side of the affected ear and noted more rapid resolution in affected right ears than affected left ears.

- Portoain-Shuhaiber and Cullinan (1984) stratified by racial/ethnic origin and noted quicker resolution of OME in African children than in Indian or Caucasian children.

- Moller and Tos (1990) found that different tympanometry instruments, which they described as impedance tympanoscopy and impedance audiometry, gave different rates of OME among the same group of children.

- Zielhuis, Rach, and van den Broek (1990) found that season and age at the end of the episode had a statistically significant effect on OME resolution rates, while gender, upper respiratory tract infection, and history of AOM did not, although they presented only percentages without denominators.

Summary

We found sufficient data only to analyze the rates of resolution of OME among children older than 3 years old. No data existed on the role of such factors as duration of OME, initial occurrence of AOM, or presence of OME in one vs. both ears. Data on the roles of gender, at-home vs. day care, race/ethnicity, time of year, age of onset, or side of ear were too limited to draw any conclusions.

Key Question 2: What Are the Effects of Early-Life OM on Long-Term Speech and Language Development?

- Do infants and preschool children with longer-duration early-life OME as compared to those with shorter duration OME have greater delays in speech and language development (receptive or expressive) later in life?

- Is OME-associated conductive hearing loss in the first 3 years of life a risk factor for speech and language developmental delays?

- What are the risk factors that interact with the effect of OME on speech and language development in infants and preschool children?

Literature Review

After secondary screening of the 449 articles retrieved for review, we identified 112 articles that fell within the scope of this question. Tertiary screening identified 20 studies that fulfilled the 5 criteria for analysis. (Table 26) The five criteria included the following:
1) OM was diagnosed/assessed before the age of 3 years, 2) speech or language outcome was measured at or before the age of 22 years, 3) a prospective cohort study design was used, 4) OM was graded, and 5) speech or language outcome was measured beyond 3 years of age. Of the 20 studies, 17 were prospective cohort studies and three (Freeark, 1992; Fischler, 1985; Paul, 1993) were retrospective-prospective studies in which OM history was retrospectively obtained, but outcome measures were prospectively obtained. We did not exclude any studies based on their quality.

When we reviewed the studies included in the 1994 OME Guideline for inclusion in our analysis (Stool, Berg, Berman et al., 1994), we excluded six of the studies. The Friel-Patti and Finitzo (1990) study was excluded, because it was not a prospective cohort study and did not report outcomes for children over 3 years of age. The Friel-Patti (1982) study was excluded because no outcomes were measured in children over 3 years of age. The studies of Lous and Fiellau-Nikolajsen (1988) and Rach, Ziehlhuis, van Baarle, and colleagues (1991 were excluded, because the investigators did not measure OM severity before the age of 3 years. The Rach, Zielhuis, and van den Broek (1988) study was excluded, because it was not a prospective cohort study. The study by Wright, Sell, McConnell, and colleagues (1988) was excluded, because no outcomes were measured in children over 3 years of age. Finally, unlike the OME Guideline (Stool, Berg, Berman et al., 1994), we included the studies by Klein (1988) and Freeark (1992), because study quality was not one of our rejection criteria.

Table 27 lists the author, year and cohort of 22 prospective cohort studies that were excluded from our analysis because they did not report outcomes for children older than 3 years of age.

These 22 studies actually included only 12 cohorts. Relevant findings for seven of these 12 cohorts were included in our assessment.

Findings

Evidence Table 2 presents the study characteristics, population characteristics, risk factors, and outcome findings for the 20 cohort studies that responded to this question. Speech and language outcomes were examined in a total of 12 cohorts of children. Of these 12 cohorts, nine included children primarily from specific populations, such as a particular ethnic or racial group or a particular socioeconomic group. It is also important to reiterate that cohorts were excluded if they consisted exclusively of children with craniofacial defects, primary mucosal disorders, immunodeficiencies, or a genetic disorder. None of the studies specifically assessed children who already had speech, language, or other developmental delays. One of the studies focused specifically on persistent bilateral OM.

For the purpose of responding to this question, we defined early life otitis media as positive otitis media history prior to 3 years of age. **Table 28** summarizes the definitions of positive or negative history of otitis media used in these studies and the diagnostic method for OM. The definition of positive or negative OM history varied among the studies. Some studies used percentage of visits during a specified time, some used number of visits with OM, some used days spent with effusion, while others combined various criteria.

The age during which the outcome was measured also varied. This age ranged from one to 3 years. The diagnostic method and the examiner also varied. Several studies based diagnosis on chart review or parents' record. A few studies used pneumatic otoscopy performed by pediatricians, otolaryngologists, or trained professionals.

The quality of the studies included in the evidence table is summarized in **Table 29**. Of the 20 studies, five (25 percent) received a score of six of a possible of eight points; four (20 percent) scored five points; six (30 percent) scored four points; four (20 percent) scored three points, and one (5 percent) scored two points.

Table 30 presents a summary of the key characteristics of the 20 cohort studies including the cohort, age of OM history, age at outcome measure, major outcome statistical analysis, outcome measure of interest, diagnostic procedure(s), and notes. The table is organized by the outcome measures for this key question, namely cognition, expressive language, receptive language, expressive speech, and receptive speech. Since multiple outcomes could be included in one study, a study may appear in several rows. The entries in the "Test" column should be interpreted with caution, because subtests versus global tests were not distinguished.

The factor we used to determine which studies to examine further by meta-analysis was the major outcome measure, which indicates the type of statistical measured used. Whenever three or more studies reported the same outcome measure, we considered pooling the data. Based on this strategy, we conducted three possible meta-analyses to derive the pooled difference between positive and negative early-life otitis media for: expressive language development, receptive language development, and development of cognitive verbal intelligence.

Expressive Language Development **Table 31** presents the findings of the three cohort studies that addressed expressive language development. All three studies measured OM history prior to 3 years of age. Although the pooled standardized mean difference showed an increase of about 14.5 percent (95% C.I.: -49.2%, 20.2%) of a standard deviation in the expressive language

measure for the group of children with no early-life OM history compared with those with a positive history, this pooled estimate is not significantly different from zero. Thus, the available data do not support the hypothesis that an OM history prior to 3 years of age has an effect on expressive language development. However, the 95% confidence intervals on our pooled results do not exclude a clinically important effect size of almost 0.5, meaning no strong conclusions can be drawn. The Chi-squared test of heterogeneity showed that the standardized difference was not significantly different among the studies. However, the ages at which outcome was measured and the tests used were not uniform across the studies.

Receptive Language Development **Table 32** presents the findings of the four cohort studies that addressed receptive language development. All four studies measured OM history at less than 3 years of age. Although the pooled standardized mean difference showed an increase of about 10.3 percent (95% C.I.: -28.9%, 49.5%) of a standard deviation of the receptive language measure in the group of children with no early-life otitis media history, this pooled estimate is not significantly different from zero. Thus, the available data do not support the hypothesis that an OM history prior to 3 years of age has an effect on receptive language development. However, the 95% confidence intervals on our pooled results do not exclude a clinically important effect size of almost 0.5, meaning no strong conclusions can be drawn. The Chi-squared test of heterogeneity showed that the standardized mean difference was not significantly different among the studies. However, the age at which outcome was measured and the test used were not uniform across the four studies. In addition, although the racial/ethnic composition was not reported precisely, one of the studies included primarily African-American children, another primarily American Indian children, and the third primarily Caucasian children in private practice.

Cognitive Verbal Intelligence **Table 33** presents the findings of the three cohort studies that addressed development of cognitive verbal intelligence. All three studies examined cognitive verbal intelligence, and all measured OM history at less than 3 years of age. Although the pooled standardized mean difference showed an increase of about 23 percent (95% C.I.: -20%, 65%) of a standard deviation in the expressive language measure in the group of children with no early life OM history, this pooled estimate is not significantly different from zero. Thus, the available data do not support the hypothesis that an OM history prior to 3 years of age has an effect on the development of cognitive verbal intelligence. However, the 95% confidence intervals on our pooled results do not exclude a clinically important effect size of almost 0.5, meaning no strong conclusions can be drawn. The Chi-squared test of heterogeneity showed that the standardized mean difference was not significantly different among studies. However, the age at which outcome was measured and the test used were not uniform across the studies. Further, two of the study populations were primarily African-American. The third study population was of lower socioeconomic status.

Summary

The data do not support an effect of early-life OME on language development or cognitive verbal intelligence. However, differences among the cohorts and study conditions and the wide

95% confidence intervals make it difficult to conclude that there is no effect. We found insufficient data to assess early-life OME on speech development.

Key Question 3: What are the Effects of Early-Life OM on Long-Term Hearing?

- Do infants and preschool children with longer duration early life OME as compared to those with shorter duration OME have permanent (or sensorineural) hearing loss later in life? One specific formulation of this question is: Is OME-associated conductive hearing loss in the first 3 years of life a risk factor for permanent (or sensorineural) hearing loss later in life?

- What are the risk factors that interact with the effect of OME on hearing later in life (unilateral or bilateral) in infants and preschool children?

Literature Review

After secondary screening of the 449 articles we retrieved for review, we identified 186 articles that fell within the scope of this question. Tertiary screening identified 12 studies that fulfilled the five criteria for analysis. The five criteria included 1) OM was diagnosed/assessed before the age of 3 years, 2) hearing outcome was measured at or before the age of 22 years, 3) a prospective cohort study design was used, 4) OM was graded, and 5) hearing outcome was measured after 3 years of age. Of the 12 cohort studies, four were excluded from further analysis. The reasons for exclusion of the four studies are presented in **Table 34**. **Table 35** lists the eight studies included in the evidence table and considered for analysis. **Table 36** lists the author, year, and cohort of 10 studies that were not included in the analysis because they failed to report findings for children over 3 years of age.

Findings

Evidence Table 3 presents the study characteristics, population characteristics, risk factors, and findings of the eight cohort studies considered for this question.

Table 37 summarizes the definitions of positive and negative history of OM used in these studies and the diagnostic method for OM. The definitions of positive and negative OM history varied from one study to another. These variations were similar to those identified in the studies of speech and language development.

The study quality of the eight studies included in **Evidence Table 3** is summarized in **Table 38**. Of the eight studies, 1 (12.5 percent) received a score of 6 of a possible of 8 points; 4 (50 percent) scored 5 points; 2 (25 percent) scored 4 points; and one (12 percent) scored 2 points.

The age at which OM history was taken, age at outcome measure, and type of outcome measure for the 8 cohort studies are displayed in **Table 39**. With the various combinations of age at outcome and type of outcome measure that characterized the studies, only one combination, percentage of conductive hearing loss at 6 to10 years of age, was considered sufficiently clinically similar to justify statistical pooling: four studies reported this outcome

measure at this age range. Three of the four studies (Fischler 1985, Harsten 1993 and Sorri 1995) reported treatment, including oral antibiotics, myringotomy, and tympanostomy tube, for OM episodes; Kaplan (1973) did not address treatment. Sorri (1995) used 20 dB as the air-conduction threshold above which hearing loss was defined. The other three studies used 25 dB as the threshold.

We conducted two meta-analyses, one including all four studies and another that excluded the Sorri 1995 studies, which used a different threshold for hearing loss from that of the other three studies. The meta-analysis findings are reported in **Table 40**. The pooled risk of conductive hearing loss among 346 children who had early-life OM was 22 percent (95% CI: 7% to 36%), compared with 6 percent (95% CI: 1% to 12%) among 237 children who did not have a history of early-life OM. The pooled difference in rate of hearing loss between those with an early-life OM history and those without was 11 percent (95% CI: 3% to 19%), and the pooled risk ratio was 2.6 (95% CI: 1.6 to 4.2). Thus, an early-life history of OM was significantly associated with conductive hearing loss. Neither the studies pooled for the rate difference nor the studies pooled for the risk ratio were statistically heterogeneous. **Figure 1** presents the shrinkage plot and **Figure 2** presents the funnel plot for the rate difference and **Figure 3** and **Figure 4** present similar plots for the risk ratio of hearing loss. These figures show that negative early-life OM history is more favorable in term of conductive hearing loss at age 6 to 10 years. The funnel plots for the risk difference and the risk ratio (**Figure 2** and **Figure 4,** respectively) showed no indication of publication bias. Neither the adjusted rank correlation tests (Begg, 1999) nor the regression asymmetry test (Egger, 1997) indicated publication bias for either statistic ($p>0.99$ for the risk difference; $p=0.31$ for the risk ratio; $p=0.71$ for the risk difference; $p=0.28$ for the risk ratio; respectively). However, these results should be interpreted with caution. Only four studies were included in the analysis. Moreover, each of these cohorts included relatively homogeneous populations of children, one from Finland, another from Sweden, one primarily of American Indian children, and another primarily of Eskimo children. The sensitivity analysis that excluded the Sorri (1995) study did not change the conclusions. However, the exclusion of the Sorri study reduced the variability of the rate difference and lowered the pooled difference in rate of conductive hearing loss between those with a positive OM history and those without from 11 percent (95% CI: 3% to 19%) to 8 percent (95% CI: 4% to 13%).

Summary

The results support that history of early-life OME is associated with increased risk for conductive hearing loss. However, the number of studies with similar outcome measures is small. We found insufficient data to assess the early-life OM on permanent hearing loss.

Key Question 4: Diagnostic Methods for OME

What are the sensitivity, specificity, and predictive values for alternative methods of diagnosing OME compared with one of the reference standards?

These methods include, but are not limited to:

- signs/symptoms
- non-pneumatic otoscopy
- pneumatic otoscopy, validated or non-validated examiner
- binaural micro-tympanoscopy
- portable tympanometry
- professional tympanometry
- quantitative tympanometry
- acoustic reflectometry (specify model and year)
- otoacoustic emissions
- audiometry, air or bone conduction thresholds.

Literature Review

After secondary and tertiary screening of the 449 articles we retrieved for review, we identified 75 articles that fell within the scope of this question. When we compared our list with the 1994 OME Guideline (Stool, Berg, Berman et al., 1994), we found five studies that were included in the 1994 OME Guideline but not in our assessment. We excluded the Kaleida (1992) and the Shurin, Pelton and Finkelstein (1977) studies, because data were not abstractable. The McDermott and Giebink, Le, and colleagues (1983) and the Teele and Teele (1984) studies were excluded, because they did not address the scope of this question. The Lampe, Weir, Spier, and colleagues (1985) study was excluded, because it was a duplicate of another study. We included three studies that were rejected by the developers of the 1994 OME Guideline, because we did not reject any studies based on study quality. These were the studies by Haughton (1977), Karma (1989), and Marchart (1986).

Of the 75 articles accepted for data abstraction, we included 52 studies in our assessment (listed in **Table 41)**. **Table 42** lists the reasons for exclusion of the 23 remaining articles. **Evidence Table 4** presents the study characteristics, the characteristics of the study population, and the study findings of each of the 52 studies included.

Of the 52 studies, 33 (63 percent) scored three points or fewer on our six-point quality scale. Of the 19 studies that scored more than three points, 15 studies scored four points, three studies scored five points, and one study scored six points. Most of the studies in this group are of poor quality, a finding that replicates those of Lijmer, Mol, Heisterkamp and colleagues (Lijmer, Mol, Heisterkamp et al., 1999).

Findings

Evidence Table 4 presents the sensitivity, specificity, positive predictive value, negative predictive value, accuracy, and prevalence of OME in the cohort for each comparison of diagnostic methods and reference standards listed within the scope of this assessment.

Table 43 summarizes the number of comparisons for each diagnostic method and reference standard pair. On the basis of these numbers we selected groups of three or more studies for meta-analysis, from which we derived pooled random effect estimate, 95% confidence intervals, and measure of heterogeneity for sensitivity, specificity, positive predictive value, negative predictive value, accuracy, and prevalence.

Tables 44 through **52** present the results of the meta-analyses that compare sensitivity, specificity, and prevalence rate for acoustic reflectometry at <=5 or >5 RU (reflective units), pneumatic otoscopy, portable tympanometry, professional tympanometry using acoustic reflex at 500 or 1000 Hz, professional tympanometry using static compensated acoustic admittance at 0.1, professional tympanometry using static compensated acoustic admittance at 0.2, professional tympanometry using static compensated acoustic admittance at 0.3, professional tympanometry using B curve as abnormal, and professional tympanometry using B or C2 curves as abnormal, respectively, using myringotomy as the reference standard. Findings excluding duplicates are summarized in **Table 53**. The receiver-operator characteristic points that correspond to sensitivity versus (1-specificity) are plotted in **Figure 5.** The receiver operator characteristic points showed that pneumatic otoscopy was closest to the optimal operating point where both sensitivity and specificity would be 100 percent.

Among the nine diagnostic methods, pneumatic otoscopy and professional tympanometry (using flat or B or C2 curve as abnormal) had the highest sensitivity at 93.8 percent (95% CI: 91.4, 96.3) and 93.8 percent (95% CI: 91.1, 96.4) compared with myringotomy, respectively. The diagnostic test with the highest specificity was professional tympanometry (using static compensated acoustic admittance at 0.1) at 95.0 percent (95% CI: 88.5, 100).

If we consider both sensitivity and (1 minus specificity) in the receiver-operator characteristic display in **Figure 5**, pneumatic otoscopy is closest to the optimal operating point where both sensitivity and specificity are 100 percent. The pooled sensitivity was 94 percent (95% CI: 91%, 96%), and the pooled specificity was 80 percent (95% CI: 75%, 86%). These findings were based on 2,694 children from 7 studies that reported a pooled prevalence of OME of 63 percent (95% CI: 58%, 67%). The estimated prevalence rates ranged from 56 percent to 71 percent, which indicated significant heterogeneity ($p<0.001$). We used the pooled sensitivity and specificity for pneumatic otoscopy and derived the positive and negative predictive values for various prevalence levels. **Figure 6** provides such a plot.

Table 54 shows an analysis of the study quality of the diagnostic tests included in the meta-analyses. With the exception of the study by Babonis (Babonis, Weir, Kelly, 1991), which scored five, all studies scored four or less out of a maximum of six possible points. The majority of the studies did not fulfill criterion four on representativeness of patient sample in clinical practice or criterion five on determination of reproducibility of test results. Among the seven studies used to derive the pooled estimates of sensitivity and specificity for pneumatic otoscopy, two studies scored the minimum of one point, two scored two points, one scored three points, and two scored four points.

In **Table 54,** we also compare the qualifications of the examiner performing the diagnostic test for each study. Of the seven studies in the comparison between pneumatic otoscopy and myringotomy, two studies did not specify the test performer, one study specified that a senior registrar and a senior house officer performed the test, and the remaining four specified that either a pediatrician or an otolaryngologist performed the test. However, whether the test performer was trained or untrained was not specified.

Summary

The meta-analyses revealed that pneumatic otoscopy and professional tympanometry had the highest sensitivity compared with myringotomy. While the diagnostic test with the highest specificity was professional tympanometry (using static compensated acoustic admittance at 0.1), pneumatic otoscopy optimized both sensitivity and specificity. However, the poor quality of many of the studies included in the analysis must be considered. Moreover, most studies failed to provide enough information to assess the qualifications of testers.

Table 17: Cohorts and Articles Relevant to Question 1 (n=40)

Cohort Identifier	Relevant Articles (ID, authors, year)	Comments
Birch	862 Birch and Elbrønd (1984)	0.75- to 7-year old children followed from 1/1982 to 4/1982.
Casselbrant I	1000 Casselbrant, Brostoff, Ashoff, and Bluestone (1985) 2929 Casselbrant, Brostoff, Ashoff, and Bluestone (1990)	2- to 5-year old children followed from 9/1981 to 8/1983.
Casselbrant II	2929 Casselbrant, Brostoff, Ashoff, and Bluestone (1990)	5- to 12-year old children followed from 9/1984 to 5/1985.
Ernston	1202 Erston and Sundberg (1984)	Children 1- to 11-years old embedded in a controlled trial.
Fiellau-Nikolajsen I	1237 Fiellau-Nikolajsen and Lous (1979) 1235 Fiellau-Nikolajsen (1979) 1242 Fiellau-Nikolajsen (1981) 3051 Lous and Fiellau-Nikolajsen (1988)[a]	Children in Hjoerring, Denmark first examined in 1976.
Fiellau-Nikolajsen II	1235 Fiellau-Nikolajsen (1979) 1245 Fiellau-Nikolajsen (1983)	Children in Hjoerring, Denmark first examined in 1978.
Fiellau-Nikolajsen III	1777 Lous and Fiellau-Nikolajsen (1981)	Children in Hirtshals and Sindal, Denmark first examined in 1978.
Holmquist	1494 Holmquist, Fadala, and Qattan (1987)	7- to 9.5-year old children followed 2/1983 to 4/1983.
Lamothe	1714 Lamothe, Boudreault, Blancette, Tetreault, and Poliquin (1981)	First graders followed over a 6 week period in 1979.
Leiberman	1735 Leiberman and Bartal (1986)	2- to 12-year old children who had a follow-up exam after a 2.5 year delay in ventilating tube placement.
Marchisio	9 Marchisio, Principi, Salpietro, Boschi, Chetri, Caramia, Longhi, Reali, Meloni, DeSantis, Sacher, and Cupido (1998)	Primary school children followed for 12 weeks after the initial exam, and then a subset were randomized to a placebo group for another 8 weeks of follow-up.
Mills	1927 Mills and Vaughan-Jone (1992)	Prospective single cohort embedded in a comparative cohort of children 1- to 14-years old who had a follow-up exam about 2 months after the initial visit.
Portoian-Shuhaiber	2184 Portoian-Shuhaiber and Cullinan (1984)	5- to 6-year old children followed for 10 weeks in 1979.
Renvall I	2240 Renvall, Lidén, Jungert, and Nilsson (1978)	10- to 11- year old children examined after a 3-year interval.
Renvall II	2242 Renvall, Anniansson, and Lidén (1982)	4-year old children followed over a 12 week period in 1980.
Reves	2243 Reves, Budgett, Miller, Wadsworth, and Haines (1985)	3- to 6-year old children followed 11/1983 to 2/1984.
Roberts	2262 Roberts, Johnson, Carlin, Turczyk, Karnutta, and Yaffee (1995)	Newborns followed for 2 months after birth.
Robinson	2270 Robinson, Allen, and Root (1988)	Infants followed for 6 weeks.

[a] The article did not include abstractable data relevant to the specific Question 1 outcome measures.

Table 17: Cohorts and Articles Relevant to Question 1 (n=40) (Continued)

Cohort Identifier	Relevant Articles	Comments
Sly I	2457 Sly, Zambie, Fernandes, and Frazer (1980)	4- to 5-year old children recruited in 2/1977.
Sly II	2457 Sly, Zambie, Fernandes, and Frazer (1980)	4- to 5-year old children recruited in 9/1977.
Tos I	1486 Holm-Jensen, Sørenson, and Tos (1981) 2636 Tos (1981) [a] 543 Tos, Holm-Jensen, Sørenson, and Mogensen (1982) 2639 Tos (1983) [a] 2642 Tos (1984) [a] 4834 Tos (1984) [a] 4835 Tos (1988) [a]	Children born in 1975 and followed from 5/1979 to 2/1985.
Tos II	2629 Tos (1979) 2190 Poulsen and Tos (1980) 2631 Tos (1980) [a] 2634 Tos (1980) 2593 Thomsen and Tos (1981) 2639 Tos (1983) [a] 2642 Tos (1984) [a] 4834 Tos (1984) [a] 4835 Tos (1988) [a]	Children born in 1976 and followed from 11/1977 to 2/1985.
Tos III	2189 Poulsen and Tos (1978) 2627 Tos (1979) 2631 Tos (1980) [a] 2634 Tos (1980) 2639 Tos (1983) [a] 2642 Tos (1984) [a] 4834 Tos (1984) [a] 4835 Tos (1988) [a]	Children born in 1977 and followed from 1-2/1977 to 2/1985.
Tos IV	1946 Moller and Tos (1990)	Children checked daily for 30 days.
van Balen	91 van Balen, De Melker, Touw-Otten (1996)	6-month to 6-year old children followed for 3 months in the early 1990's.
Williamson	2791 Williamson (1994)	5- to 8-year old children followed from 1988-1989 to 1991.
Wilmot	2795 Wilmot (1988)	6-month to 10-year old children were followed for 12 months after developing OME after AOM.
Zielhuis	2863 Zielhuis, Rach, and van den Broek (1990)	2- to 4-years old children followed from 1982-1983.

[a] The article did not include abstractable data relevant to the specific Question 1 outcome measures.

Table 18: Disposition of Articles From Rosenfeld's Natural History Assessment

ID Number	Author	Year	Included	Excluded	Reason for Exclusion From Assessment
1000	Casselbrant	1985	X		
1238, 1240	Fiellau-Nikolajsen	1980		X	1238 is a study of tympanometry as a diagnostic tool. 1240 is the same cohort as in 1245 and does not present sufficient data tracking individual cases or episodes of OME.
2242	Renvall	1982	X		
2593	Thomsen	1981	X		2593 uses same cohort as 2190, 2629,2631, 2639, 2642, 4834, and 4835.
2627	Tos	1979	X		
2634	Tos	1980	X		
543	Tos	1982	X		543 uses same cohort as 1486, 2636, 2639,2642, 4834, and 4835.
2791	Williamson	1994	X		
2863	Zielhuis	1990	X		
2857	Zeisel	1995		X	2857 included 13% with purulent OME. Antibiotic administration and other interventions for either purulent or non-purulent OME. Insufficient data track individual cases or episodes of OME.
2243	Reves	1985	X		
91	van Balen	1996	X		Randomized controlled trial with an initial 3-month watchful waiting period of children with OME.
1777	Lous	1981	X		

Table 18: Disposition of Articles From Rosenfeld's Natural History Assessment

ID Number	Author	Year	Included	Excluded	Reason for Exclusion From Assessment
1000	Casselbrant	1985	X		
1238, 1240	Fiellau-Nikolajsen	1980		X	1238 is a study of tympanometry as a diagnostic tool. 1240 is the same cohort as in 1245 and does not present sufficient data tracking individual cases or episodes of OME.
2242	Renvall	1982	X		
2593	Thomsen	1981	X		2593 uses same cohort as 2190, 2629,2631, 2639, 2642, 4834, and 4835.
2627	Tos	1979	X		
2634	Tos	1980	X		
543	Tos	1982	X		543 uses same cohort as 1486, 2636, 2639,2642, 4834, and 4835.
2791	Williamson	1994	X		
2863	Zielhuis	1990	X		
2857	Zeisel	1995		X	2857 included 13% with purulent OME. Antibiotic administration and other interventions for either purulent or non-purulent OME. Insufficient data track individual cases or episodes of OME.
2243	Reves	1985	X		
91	van Balen	1996	X		Randomized controlled trial with an initial 3-month watchful waiting period of children with OME.
1777	Lous	1981	X		

Table 19 Question 1: Articles Excluded During Data Abstraction

ID#	Author	Year	Reason Not Included
705	Aniansson	1985	Screening study over 2 years. Some retest of same subjects, but no control over treatment.
3051	Lous	1988	Data not abstractable as presented. Available counts are in essential agreement with articles 1235, 1237, and 1242 apart from minor differences in number of tympanogram types, e.g. type A in Jan 1975 (629 in articles 1235, 1237, and 3051 and 631 in article 1242) and types B, C1, and C2 in Jun-July 1976 (32, 31,, and 42 respectively in articles 1235, 1242, and 3051 and 37, 28, and 40 respectively in article 1237).
2578	Teele	1980	Data not abstractable; data on persistent OME after AOM are not presented and cannot be derived.
2631	Tos	1980	Article does not present any new information other than the actual initial counts of tympanogram types for the 1976 and 1977 Tos cohorts utilized in the evidence tables for article 2629 and 2627.
2636	Tos	1981	Data not abstractable as presented relative to the data on 3- and 6-month follow-up presented in article 1486. Individual cases of type B, C1, or C2 tympanograms cannot be tracked from 2/1979 to 11/1979 or 2/1980.
2639	Tos	1983	The only relevant new information is that the authors mention that in the 1976 cohort, 50% of tympanogram type B changed to types A or C over the first 3 months of the study. However, the number with type B who presented for the 3-month follow-up exam is not given so per cent change cannot be calculated.
2642	Tos	1984	Article does not present any abstractable data relevant to Q1.
4834	Tos	1984	Article does not present any new data relevant to Q1.
4835	Tos	1988	Article does not present any new data relevant to Q1.
2795	Wilmot	1988	Article presented data on OME following AOM which was eliminated as a condition of interest because it had been studied in a recent evidence analysis.

Table 19 Question 1: Articles Excluded During Data Abstraction

ID#	Author	Year	Reason Not Included
705	Aniansson	1985	Screening study over 2 years. Some retest of same subjects, but no control over treatment.
3051	Lous	1988	Data not abstractable as presented. Available counts are in essential agreement with articles 1235, 1237, and 1242 apart from minor differences in number of tympanogram types, e.g. type A in Jan 1975 (629 in articles 1235, 1237, and 3051 and 631 in article 1242) and types B, C1, and C2 in Jun-July 1976 (32, 31,, and 42 respectively in articles 1235, 1242, and 3051 and 37, 28, and 40 respectively in article 1237).
2578	Teele	1980	Data not abstractable; data on persistent OME after AOM are not presented and cannot be derived.
2631	Tos	1980	Article does not present any new information other than the actual initial counts of tympanogram types for the 1976 and 1977 Tos cohorts utilized in the evidence tables for article 2629 and 2627.
2636	Tos	1981	Data not abstractable as presented relative to the data on 3- and 6-month follow-up presented in article 1486. Individual cases of type B, C1, or C2 tympanograms cannot be tracked from 2/1979 to 11/1979 or 2/1980.
2639	Tos	1983	The only relevant new information is that the authors mention that in the 1976 cohort, 50% of tympanogram type B changed to types A or C over the first 3 months of the study. However, the number with type B who presented for the 3-month follow-up exam is not given so per cent change cannot be calculated.
2642	Tos	1984	Article does not present any abstractable data relevant to Q1.
4834	Tos	1984	Article does not present any new data relevant to Q1.
4835	Tos	1988	Article does not present any new data relevant to Q1.
2795	Wilmot	1988	Article presented data on OME following AOM which was eliminated as a condition of interest because it had been studied in a recent evidence analysis.

Table 20 Study Quality for Studies Included in Evidence Table on Natural History

ID	Author	Year	Study Quality Score[a]
862	Birch	1984	3 (1,1,1,0,0,0)
1000	Casselbrant	1985	3 (1,1,1,0,0,0)
2929	Casselbrant	1990	3 (1,1,1,0,0,0)
1202	Ernstson	1984	2 (1,0,0,1,0,0)
1235	Fiellau-Nikolajsen	1979	4 (1,1,1,0,0,1)
1237	Fiellau-Nikolajsen	1979	4 (1,1,1,0,0,1)
1242	Fiellau-Nikolajsen	1981	4 (1,1,1,0,0,1)
1245	Fiellau-Nikolajsen	1983	3 (1,1,1,0,0,0)
1486	Holm-Jensen	1981	3 (1,1,1,0,0,0)
1494	Holmquist	1987	3 (1,1,1,0,0,0)
1714	Lamothe	1981	4 (1,1,1,1,0,0)
1735	Leiberman	1986	2 (1,1,0,0,0,0)
1777	Lous	1981	3 (1,1,1,0,0,0)
9	Marchisio	1998	3 (1,1,1,0,0,0)
1927	Mills	1992	1 (1,0,0,0,0,0)
1946	Moller	1990	3 (1,1,1,0,0,0)
2184	Portoian-Shuhaiber	1984	3 (1,1,1,0,0,0)
2189	Poulsen	1978	3 (1,1,1,0,0,0)
2190	Poulsen	1980	3 (1,1,1,0,0,0)
2240	Renvall	1978	3 (1,1,1,0,0,0)
2242	Renvall	1952	4 (1,1,1,1,0,0)
2243	Reves	1985	3 (1,1,1,0,0,0)
2262	Roberts	1995	4 (1,1,1,0,1,0)
2270	Robinson	1988	1 (1,1,0,0,0,0)
2457	Sly	1980	3 (1,1,1,0,0,0)
2593	Thomsen	1981	3 (1,1,1,0,0,0)
2627	Tos	1979	3 (1,1,1,0,0,0)
2629	Tos	1979	3 (1,1,1,0,0,0)
2634	Tos	1980	3 (1,1,1,0,0,0)
543	Tos	1982	3 (1,1,1,0,0,0)
91	van Balen	1996	3 (1,1,1,0,0,0)
2791	Williamson	1994	2 (1,1,0,0,0,0)
2863	Zielhuis	1990	3 (1,1,1,0,0,0)

[a] The six components of study quality are: a prospective cohort study; outcome clearly defined; time point at which outcome measured clearly defined; subjects followed without any intervention; blinded assessment of outcome; and point and variability estimates provided for main outcome measures. 1 indicates presence and 0 indicates absence

Table 21 Tympanometry Definitions in Natural History Cohorts Utilizing Tympanometry as the Sole Diagnostic Method

Tympanogram Type	A		As		B		
Cohort[a]	Pressure (mmH_2O)	immitance	Pressure (mmH_2O)	immitance	Pressure (mmH_2O)	immitance	Comments
Birch	> -100					< 0.25ml[c,1]	[1]stapedial reflex absent and max compliance unreadable; Madsen Electronics tympanoscope, model ZS 330, 226 Hz probe tone
Fiellau-Nikolajsen I	> -100	> 0.1 [d]			200 to -400[1]	≤ 0.1[d]	[1]or indeterminable
Fiellau-Nikolajsen II							middle-ear effusion=flat curve or <= -100 with absent middle ear reflexes
Fiellau-Nikolajsen III	> -100					multiple criteria[1]	[1]type B=otoadmittance < 0.20millimhos, absolute gradient < 0.04millimhos and absence of ipsilateral acoustic reflex; Grason-Stadler Middle Ear Analyzer 1722
Holmquist	50 to -99				flat curve		Madsen ZA 330, 226 Hz probe tone
Portoian-Shuhaiber							abnormal defined as an abnormal tympanometric curve and/or absent acoustic reflex; Grason-Stadler Middle-ear Analyser (Model 1722)
Tos I	0 to -99				flat curve		Madsen ZO 70 tympanometer, 220 Hz probe tone
Tos II	0 to -99				flat curve		Madsen ZO 70 tympanometer, 220 Hz probe tone
Tos III	>-100				flat curve[1]		[1]<= 0.1[e]; Madsen ZO 70 impedance meter
Tos IV AZ7	99 to –99				flat curve without impedance minimum[1]		[1]or with a measurable impedance minimum and relative gradient < 0.1; Impedance audiometer AZ 7 (Interacoustics)
ZS 331	99 to –99				flat training[1]		[1]or compliance below 0.25ml[c] and absent ipsilateral stapedial reflex; Impedance tympanoscope ZS 331 (Madsen Electronics)
Reves	-100 to 50	>0.3[b]			< -100	low[b]	tympanometer 85 AR 11 (American Electro Medics)
Robinson	-149 to +50	>0.2ml[c]					types As, B, C, and Cs are failures; Maico MA 610 portable impedance screener, 226 Hz probe tone
Sly I and II					flat curve[1]		[1]or compliance < 0.3ml[c] or peak compliance occurred at or below -100 mmH_20; Teledyne Avionics acoustic impedance meter model TA-1D
van Balen	-99 to 200	≥ 0.2mmho[c]			≤ -400	< 0.2mmho[c]	
Williamson	200 to -99				flat curve		Grayson-Stadler [sic] Earscan impedance audiometer
Zielhuis	≥ -99	≥ 0.2ml[c]			≤ -400	< 0.2ml[c]	Grason-Stadler-model 27

[a]see Table 17 for associated cohort articles [b]impedance [c]admittance [d]relative gradient [e]impedance slope

Table 21 Tympanogram Definitions in Natural History Cohorts Utilizing Tympanometry as the Sole Diagnostic Method (continued)

Tympanogram Type	C		C1		C2		
Cohort[a]	Pressure (mmH_2O)	immittance	Pressure (mmH_2O)	immittance	Pressure (mmH_2O)	immittance	Comments
Birch	≤ -100						
Fiellau-Nikolajsen I			-100 to -199	>0.1[d]	-200 to -400	>0.1[d]	
Fiellau-Nikolajsen II							middle-ear effusion=flat curve or <= -100 with absent middle ear reflexes
Fiellau-Nikolajsen III	≤ -100						
Holmquist	100 to -300						
Portoian-Shuhaiber							abnormal defined as an abnormal tympanometric curve and/or absent acoustic reflex
Tos I			-100 to -199		-200 to -350		
Tos II			-100 to -199		-200 to -350		
Tos III	-100 to -300	>0.1[e]	-100 to -199		-200 to -350		
Tos IV AZ 7			-100 to –199		>-200		
ZS 331			-100 to –199		>-200		
Reves	< -100	>0.3[b]					
Robinson							types As, B, C, and Cs are failures
Sly I and II							
van Balen			-199 to -100	≥ 0.2mmho[c]	-399 to -200	≥ 0.2mmho[c]	
Williamson			-100 to -199		-200 to -400		
Zielhuis			-100 to -199	≥ 0.2ml[c]	–200 to –399	≥ 0.2ml[c]	

[a]see Table 17 for associated cohort articles [b]impedance [c]admittance [d]relative gradient [e]impedance slope

Table 22: OME Resolution by Ears on Newly Diagnosed OME of Unknown Duration

<6-month old cohorts

Cohort ID	Diagnostic Method[b]	antibiotic[c]	surgery[c]	Ears Resolved Interval[a]				Article(s)
				<2wk	<2m	<3m	<6m	
Roberts[e]	oto	unknown	unknown	22/24 (92%)	24/24 (100%)			2262 Roberts 1995
Tos III	Tymp	unknown	unknown			1/4[d] (25%)	1/4[d] (25%)	2627 Tos 1979

[a] interval calculated from cohort inception and not cumulative, unless otherwise noted
[b] oto=otoscopy, tymp=tympanometry (type B to A transition)
[c] Did any of the patients receive antibiotic or surgery?
[d] interval started at 6-month follow-up
[e] cumulative resolution rate

6-month to 3-year old cohorts

Cohort ID	Diagnostic Method[b]	antibiotic[c]	surgery[c]	Ears Resolved Interval[a]					Article(s)
				<6wk	<3m	<6m	<9m	<24m	
Robinson	tymp	unknown	unknown	10 of 25[d] (40.0%)					2270 Robinson 1988
Tos II	tymp	unknown	yes		6/51[1] (12%)	15/59[2] (25%)	19/51[3] (37%)	9/48[4] (19%)	[1] 2634 Tos 1980; [2] 2190 Poulson 1980; [3] 2629 Tos 1979; [4] 2593 Thomsen 1981
Tos II[e]	tymp	unknown	yes		6/51[1] (12%)	16/51[1] (31%)	24/51[1] (47%)		[1] 2634 Tos 1980

[a] interval calculated from cohort inception and not cumulative, unless otherwise noted
[b] oto=otoscopy, tymp=tympanometry (type B to A transition)
[c] Did any of the patients receive antibiotic or surgery?
[d] interval is minimum of 6wk so may be greater
[e] cumulative resolution rate

Table 22 (Continued)

>3-year old cohorts

Cohort ID	Diagnostic Method[b]	antibiotic[c]	surgery[c]	Ears Resolved Interval[a]										Article(s)
				<2wk	<3wk	<1m	<6wk	<3m	<4m	<6m	<8m	<1y	<3y	
Fiellau-Nikolajsen I[d]	tymp	unknown	yes			14/94[1] (15%)		22/91[1] (24%)		32/69[1] (46%)			33/65[2] (51%)	[1]1237 Fiellau-Nikolajsen 1979; [2]1242 Fiellau-Nikolajsen 1981
Fiellau-Nikolajsen II[d]	tymp	unknown	unknown			7/64 (11%)		16/62 (26%)						1235 Fiellau-Nikolajsen 1979
Holmquist	tymp	unknown	unknown					251/511 (49%)						1494 Holmquist 1987
Lamothe	pneum oto	no	no		24/64 (38%)		25/53 (47%)							1714 Lamothe 1981
Lamothe[d]	pneum oto	no	no		24/64 (38%)		38/53 (72%)							1714 Lamothe 1981
Renvall I	tymp	unknown	unknown										282/335 (84%)	2240 Renvall 1978
Renvall II[d]	tymp	no	no				10/40 (25%)	16/40 (40%)						2242 Renvall 1982
Sly	tymp	no	no	1/9 (11%)		4/9 (44%)	6/9 (67%)							2457 Sly 1980
Sly II	tymp	no	no	0/5 (0%)		0/5 (0%)	0/5 (0%)							2457 Sly 1980
Tos I	tymp	unknown	unknown					3/92[1] (3%); 3/87[2] (4%)		14/93[1] (15%); 14/87[2] (16%)				[1]1486 Holm-Jensen 1981; [2]543 Tos 1982
Tos I[d]	tymp	unknown	unknown					3/87[1] (3%)		17/87[1] (20%)				[1]543 Tos 1982
Williamson	tymp	unknown	yes						35/67 (52%)		52/67 (78%)	61/67 (91%)		2791 Williamson 1994

[a] interval calculated from cohort inception and not cumulative, unless otherwise noted
[b] oto=otoscopy, pneum oto=pneumatic otoscopy, tymp=tympanometry (type B to A transition)
[c] Did any of the patients receive antibiotic or surgery?
[d] cumulative resolution rates

Table 22 (Continued)

Age not stratifiable

Cohort ID	Diagnostic Method[b]	antibiotic[c]	surgery[c]	Ears Resolved Interval[a]							Article(s)
				<1m	<2m	<3m	<4m	<5m	<6m	<3y	
Casselbrant I[d]	algorithm	yes	yes	92/137 (67%)	109/137 (80%)	130/137 (95%)	134/137 (98%)	136/137 (99%)	137/137 (100%)		1000 Casselbrant 1985
Renvall I	tymp	unknown	unknown							282/335 (84%)	2240 Renvall 1978

[a] interval calculated from cohort inception and not cumulative, unless otherwise noted
[b] algorithm=algorithm based on pneumatic otoscopy, tympanometry, and acoustic reflex, tymp=tympanometry (type B to A transition)
[c] Did any of the patients receive antibiotic or surgery?
[d] cumulative resolution rate

Table 23: OME Resolution by Child on Newly Diagnosed OME of Unknown Duration

>3-year old cohorts

Cohort ID	Diagnostic Method[b]	antibiotic[c]	surgery[c]	Resolved interval[a]									Article(s)
				<2wk	<1m	<6wk	<10wk	<3m	<4m	<6m	<8m	<1y	
Fiellau-Nikolajsen II	tymp	unknown	unknown		28/81 (35%)			46/80 (58%)		53/78 (68%)			1245 Fiellau-Nikolajsen 1983
Marchisio	pneum oto, tymp	unknown	no					325/451 (72%)					9 Marchisio 1998
Portoian-Shuhaiber	tymp	unknown	unknown				65/130 (50%)						2184 Portoian-Shuhaiber 1984
Sly I	tymp	no	no	1/7 (14%)	3/7 (43%)	5/7 (71%)							2457 Sly 1980
Sly II	tymp	no	no	0/3 (0%)	0/3 (0%)	0/3 (0%)							2457 Sly 1980
Williamson	tymp	unknown	yes						22/50 (44%)		38/50 (76%)	45/50 90%)	

Age not stratifiable

Cohort ID	Diagnostic Method[b]	antibiotic[c]	surgery[c]	Resolved interval[a]		Article(s)
				<2m	<3m	
Mills	pnem oto, tymp	unknown	unknown	57/192 (30%)		1927 Mills 1992
Reves	tymp	unknown	unknown		40/68 (59%)	2243 Reves 1985
van Balen	tymp	unknown	unknown		223/443 (50%)	91 van Balen 1996

[a]interval calculated from cohort inception unless otherwise noted
[b]oto=otoscopy, tymp=tympanometry (type B to A transition)
[c]Did any of the patients receive antibiotic or surgery?

Table 24: Meta-Analysis for <6 Weeks Resolution Rate for Newly Diagnosed OME of Unknown Duration In Children Older Than 3 Years of Age

ArticleID	Author	Criterion	Age at diagnosis	Antibiotic used?	Surgery performed?	Followup interval	Number ears resolved	Total number ears	Resoluation rate in %	Random Effects Pooled Estimate (95% CI)	Test of Heterogeity statistic (P-value) Q
2457	Sly-1980	B or C to A	5yr	no	unknown	<6wk	18	32	56.3		
2457	Sly-1980	B or C to A	5yr	no	unknown	<6wk	11	22	50.0		
1714	Lamothe-1981	Otoscopy	6yr	no	unknown	<6wk	25	53	47.2		
	Total						54	107	44.9	42.3 (24.1, 60.6)	7.85 (p=0.02)

ArticleID	Author	Criterion	Age at diagnosis	Antibiotic used?	Surgery performed?	Followup interval	Number ears resolved	Total number ears	Resoluation rate in %	Random Effects Pooled Estimate (95% CI)	Test of Heterogeity statistic (P-value) Q
2457	Sly-1980	B to A	5yr	no	unknown	<6wk	6	9	66.7		
2457	Sly-1980	B to A	5yr	no	unknown	<6wk	0	5	0.0		
1714	Lamothe-1981	Otoscopy	6yr	no	unknown	<6wk	25	53	47.2		
	Total						31	67	46.3	37.2 (1.8, 72.5)	16.4 (p<0.001)

Note: Lamothe's study used otoscopy and is included in all meta-analyses

Table 25: Meta-Analysis for <3 Months Cumulative Resolution Rate for Newly Diagnosed OME of Unknown Duration In Children Older Than 3 Years of Age

ArticleID	Author	Criterion	Age at diagnosis	Antibiotic used?	Surgery performed?	Followup interval	Number ears resolved	Total number ears	Resoluation rate in %	Random Effects Pooled Estimate (95% CI)	Test of Heterogeity Q statistic (P-value)
1237	Fiellau-Nikolajsen-1979	B or C to A	3-4yr	unknown	unknown	<3mo	154	348	44.3		
1235	Fiellau-Nikolajsen-1979	B or C to A	3-4yr	unknown	unknown	<3mo	83	200	41.5		
543	Tos-1982	B or C to A	4yr	unknown	unknown	<3mo	103	393	26.2		
2242	Renvall-1982	otoscopy	4yr	no	unknown	<12wk	86	144	59.7		
	Total						426	1085	39.3	42.7 (29.3,56.1)	63.01 (p<0.001)

ArticleID	Author	Criterion	Age at diagnosis	Antibiotic used?	Surgery performed?	Followup interval	Number ears resolved	Total number ears	Resoluation rate in %	Random Effects Pooled Estimate (95% CI)	Test of Heterogeity Q statistic (P-value)
1237	Fiellau-Nikolajsen-1979	B to A	3-4yr	unknown	unknown	<3mo	22	91	24.2		
1235	Fiellau-Nikolajsen-1979	B to A	3-4yr	unknown	unknown	<3mo	16	62	25.8		
543	Tos-1982	B to A	4yr	unknown	unknown	<3mo	3	87	3.4		
2242	Renvall-1982	otoscopy	4yr	no	unknown	<12wk	16	40	40.0		
	Total						57	280	20.4	22.5 (5.9,39.0)	44.28 (p<0.001)

Table 26: List of Cohort Studies Included for Question 2

ID#	Author	Year	Cohort
1623	Kaplan	1973	Eskimo villages in the Yukon and Kuskokwim River Delta areas of Southwestern Alaska
1255	Fischler	1985	Four Indian reservations in Arizona
4657	Roberts	1986	Frank Porter Graham Child Development Center, Chapel Hill, NC
3118	Roberts	1988	Frank Porter Graham Child Development Center, Chapel Hill, NC
4806	Roberts	1988	Frank Porter Graham Child Development Center, Chapel Hill, NC
4656	Roberts	1989	Frank Porter Graham Child Development Center, Chapel Hill, NC
3117	Roberts	1991	Frank Porter Graham Child Development Center, Chapel Hill, NC
4319	Roberts	1995	Frank Porter Graham Child Development Center, Chapel Hill, NC
1373	Gravel	1992	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
4728	Gravel	1996	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
1941	Mody	1999	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
2295	Ruben	1997	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
1219	Feagans	1987	Medical and Day Care Intervention Project in Pennsylvania
2135	Paul	1993	Portland Language Development Project (PLDP), Oregon
4651	Klein	1988	The Greater Boston Otitis Media Study Group, MA
2583	Teele	1990	The Greater Boston Otitis Media Study Group, MA
1435	Harsten	1993	University Hospital of Lund, Sweden
877	Black	1993	University of Maryland Medical System, Baltimore, MD
4675	Owen	1996	University of Texas Medical Branch, TX
1277	Freeark	1992	University-based pediatric clinic in Michigan

Table 27: List of Cohort Studies Not Included for Question 2 (For Reason of Not Reporting Findings Beyond 3 Years of Age)

ID#	Author	Year	Cohort
3119	Roberts	1995	Frank Porter Graham Child Development Center, Chapel Hill, NC
4841	Wallace	1988	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
2739	Wallace	1988	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
2740	Wallace	1988	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
4842	Wallace	1992	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
667	Abraham	1996	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
4796	Petinou	1996	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
2742	Wallace	1996	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
3096	Petinou	1999	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
4671	Luloff	1993	Longitudinal study investigating the efficacy of drug prophylaxis on otitis media in greater Boston area, MA
4673	Tsushima	1993	Longitudinal study investigating the efficacy of drug prophylaxis on otitis media in greater Boston area, MA
4674	Wendler-Shaw	1993	Longitudinal study investigating the efficacy of drug prophylaxis on otitis media in greater Boston area, MA
875	Black	1988	Maryland Otitis Media Study Group, Baltimore
4708	Downs	1988	Not specified
2719	Vernon-Feagans	1996	Ongoing study of health and day-care in a semi-rural area of northeastern United States
1288	Friel-Patti	1982	Parkland Memorial Hospital, Dallas, TX
4713	Feldman	1996	Pittsburgh-area Child Development/Otitis Media Study Group, PA
4642	Paradise	2000	Pittsburgh-area Child Development/Otitis Media Study Group, PA
2819	Wright	1988	Pneumococcal vaccine study in Nashville, TN
1677	Knishkowy	1991	PROD (Promotion of Growth and Development) Program, Western Jerusalem
2579	Teele	1984	The Greater Boston Otitis Media Study Group, MA
4664	Feagans	1994	Three day-care facilities in central Pennsylvania, PA

Table 28 Definition of Early Life Positive or Negative OM History in Assessing Long-term Speech and Language Development

ID#	Author Year	Definition of Positive/Negative OM History	Definition of OM	OM Diagnosis Method
877	Black 1993	OM History defined by number of episodes of OM within the first year of life documented by otologic examination. • Positive OM History: At least 2 episodes of OM within the first year. A child could receive credit for only one bout of OM within each 29-day period. • Negative OM History: Had not experienced otitis media during first year of life	Not provided	• Based on otologic examination • Examiner not provided
1219	Feagans 1987	OM History described by frequency and duration of OM from 0 to 3 years. They were treated as continuous variables for correlation with outcome. No grouping of children by positive or negative history was done. • Frequency was calculated by counting the number of different episodes • Duration was calculated by counting the total number of days the child had effusion during the first 36 months of life.	Not provided.	• Based on pneumatic otoscopy. Beginning 1978 tympanometry was used to corroborate the diagnosis • OM diagnosed by two pediatricians and two nurse practitioners

Table 28 (Continued)

ID#	Author Year	Definition of Positive/Negative OM History	Definition of OM	OM Diagnosis Method
1255	Fischler 1985	OM History defined by number of OM episodes by age 2 years and over age 2. • The study defined three groups of children by OM history by number of episodes before and after age 2. • For our assessment, we used groups 1 and 2 as positive OM history, i.e. (>=2 attacks by age 2) and group 3 as negative OM history (<2 attacks by age 2).	Any mention of one of the following: • Acute suppurative OM: history of ear pain or fussiness with or without fever or ear drainage (less than 5 days), and physical evidence of redness with or without immobility, bulging, or a small perforation of the TM. • Serous OM: history of ear fullness, popping, or hearing loss, or an asymptomatic history; and physical evidence of TM retraction and/or immobility, with or without gray or yellow color or bubbles behind the TM • Chronic OM: history of ear drainage and perforation present for more than two weeks; and physical evidence of perforation.	• By medical record review of documented physician's clinical diagnosis.
1277	Freeark 1992	OM history severity defined by a) number of separate episodes of OM and b) total number of days of effusion over the first 3 years of life. • High OM: above median of OM severity • Low OM: below median of OM severity	Not specified	• By whom, not specified; How diagnosed, not specified. (OM history obtained from medical records).
1373	Gravel 1992	OM groups were defined by otoscopic histories during the first year of life. • OM positive: when bilateral OM was detected at 30% or more of the baby's first year visits. • OM negative: when middle ear status was rated as normal in both ears during 80% or more of the first year visits.	Not specified	• Pediatric nurse practitioners completed pneumo-otoscopic examinations during each scheduled well-baby visit

Table 28 (Continued)

ID#	Author Year	Definition of Positive/Negative OM History	Definition of OM	OM Diagnosis Method
4728	Gravel 1996	Same as 1373	Same as 1373	Same as 1373
1435	Harsten 1993	OM groups were defined by the number of AOM episodes during the first 3 years of life. • OM positive: developed six or more episodes of AOM during a 12-month period. • OM negative: no AOM episodes and less than six other acute respiratory tract infections.	• AOM was defined as an acute episode of earache in a child with red bulging eardrum(s) or purulent discharge, occasionally febrile and with signs of upper respiratory tract infection.	• By otomicropscopy performed by an otolaryngologist
1623	Kaplan 1973	OM groups were based on age of onset of first episode of otorrhea. • The study defined 3 groups: group 1-onset of first otorrhea episode during first year of life; group 2-onset of first otorrhea episode at 2-10 years, and group 3-no history of OM • For our assessment, we used group 1 as positive OM history and group 3 as negative OM history.	• Used only episodes of OM with otorrhea.	• A research nurse visited the cohort children and obtained information concerning middle ear abnormality and upper respiratory tract illness and reviewed medical records for status between visits.
4651	Klein 1988	OM history was measured by time spent with effusion during the first 2 years of life and used a 'window' of 23 days to each observation of effusion, whether accompanied by signs of illness or not. It could be shortened or extended by multiple examinations. • Group 1: time spent with effusion <32 days during first 2 years of life. • Group 2: time spent with effusion between 33-108 days during first 2 years of life. • Group 3: time spent with effusion >108 days.	• Criteria for effusion: otorrhea, gas-liquid levels visible on otoscopy or marked reduction of mobility. Tympanometric criterion: type B curve.	• By pediatricians using pneumatic otoscopy until age 3 and both pneumatic otoscopy and tympanometry in years 4 through 7.

Table 28 (Continued)

ID#	Author Year	Definition of Positive/Negative OM History	Definition of OM	OM Diagnosis Method
1941	Mody 1999	OM history defined by pneumatic otoscopy findings during first year of life • OM positive: children who had 30% or more of the 13 first-year visits with OM bilaterally • OM negative: children who had 80% or more of the 13 first-year visits with normal middle ear findings bilaterally.	• Used a 9-item otoscopic checklist to determine "clear", "suspicious," or "positive" for OM	• By trained and validated pediatric nurse practitioner using a pneumatic otoscope under the supervision of a pediatric otolaryngologist. • The PNP recorded a description of TM characteristics for each ear, using a 9-item otoscopic checklist and made the determination of "clear," "suspicious," or "positive" for OM.
4675	Owen 1996	OME history was measured by days or duration with OME durng the first 3 years of life. Middle ear status was monitored by home visits every 2 to 4 weeks, irrespective of symptoms for the first 3 years of life. • At each visit, each ear received a diagnosis of normal or OME. If two consecutive visits showed OME, the intervening days were counted as days with OME. If one visit showed OME and the next normal status, or vice versa, half of the intervening days were counted as days with OME. OME duration was defined as the proportion time a subject spent with OME (total OME days divided by total days) in the period examined. • Subjects who experienced 6 continuous weeks of OME in the first year of life were identified as at high risk.	• OME diagnosis was based on type B tympanogram or >= 5 acoustic reflectivity or visible purulent otorrhea without an otoscope.	• By trained technicians using automated screening tympanometers with a 26 Hz probe tone. Acoustic reflectivity was also measured using acoustic otoscope at 30% of visits.

Table 28 (Continued)

ID#	Author Year	Definition of Positive/Negative OM History	Definition of OM	OM Diagnosis Method
2135	Paul 1993	OM history was measured by middle ear involvement defined as either the placement of myringotomy tubes or the presence of six or more ear infections treated by a physician before the second birthday by parent report. • Positive OM: had middle ear involvement • Negative OM: lack of such middle ear involvement	• Based on parental reporting	• Based on parental reporting
4657	Roberts 1986	OM history was based on total OME duration in days during first 3 years of life. Duration of each episode of unilateral and bilateral was calculated by subtracting the data of onset of OME from the resolution date. Days of total OME was analyzed both as a continuous and categorical variable. • Group 1: days with total OME representing the lower third of the subjects • Group 2: days with total OME representing the middle third of the subjects • Group 3: days with total OME representing the upper third of the subjects.	• When middle ear fluid was seen or when the mobility of the tympanic membrane was markedly reduced or absent, OME was diagnosed. • Type B tympanograms with a flat or gradually rising shape were considered indicative of OME. • Type C tympanograms showing a maximum compliance of less than –100 mm H_2O were considered indicative of negative middle ear pressure and of an increase likelihood of effusion	• By pediatricians and pediatric nurse practitioners based on pneumatic otoscopy. 60% of the time tympanometry was used to corroborate the OME diagnosis.
4656	Roberts 1989	Same as 4657	Same as 4657	Same as 4657
3118, 4806	Roberts 1988	Same as 4657	Same as 4657	Same as 4657
3117	Roberts 1991	Same as 4657	Same as 4657	Same as 4657
4318	Roberts 1995	Same as 4657	Same as 4657	Same as 4657

Table 28 (Continued)

ID#	Author Year	Definition of Positive/Negative OM History	Definition of OM	OM Diagnosis Method
2295	Ruben 1997	OM history was based on the findings of pneumatic otoscopy at every scheduled and sick visit during the first year of life • The OM – group were those who were bilaterally free of OM at 80% or more of their first year visits and had no more than one episode of OM during the first year. • The OM+ group had bilateral OM at 30% or more of their first year visits, and had from 2 to 6 episodes during the first year.	Not specified	• Pneumatic otoscopy by a trained and validated pediatric nurse practitioner.
2583	Teele 1990	OM history was measured by the number of days with middle ear effusion (MEE) by age 3 years. Unless documented to be shorter, each episode of MEE lasted 29 days.	• Diagnosis of MEE required either visualization of a gas-liquid mixture through an intact tympanic memberane, otorrhea, or marked reduction in mobility of the TM to both positive and negative pressure. • Children whose TM(s) showed reduced mobility in response to positive pressure and normal mobility to negative pressure were considered to have only subatmospheric middle ear pressure. • The criteria for effusion using tympanometric devices included a tracing that showed no peak or a tracing that sagged below the baseline (for model 1720B)	• In private practice, three board-certified pediatricians performed 81% of all exams; in urban health center three board-certified pediatricians performed 66% of all exams. • Each center used otoadmittance meters, initially a Grason-Stadler model 1720B. Later, for children 4-7 years, used a Grason-Stadler model 1722. • To resolve ambiguous diagnoses, otoadmittance was used sporadically at 0-3 years of age, but frequently at 4-7 years.

Table 29 Study Quality for Studies Included in Evidence Table on Speech and Language Development

ID	Author(s)	Year	Study Quality Score [a]
877	Black	1993	3 (0,0,0,0,1,1,0,1)
1219	Feagans	1987	5 (1,1,1,0,0,1,0,1)
1255	Fischler	1985	5 (1,1,0,0,1,1,1,0)
1277	Freeark	1992	4 (1,0,0,0,0,1,1,1)
1373	Gravel	1992	5 (1,1,0,0,1,1,1,0)
4728	Gravel	1996	3 (0,1,0,0,0,1,1,0)
1435	Harsten	1993	6 (1,1,1,0,1,1,1,0)
1623	Kaplan	1973	6 (1,1,0,0,1,1,1,1)
4651	Klein	1988	4 (1,0,1,0,0,1,0,1)
1941	Mody	1999	4 (1,1,0,0,0,1,0,1)
4675	Owen	1996	4 (1,1,0,0,0,1,0,1)
2135	Paul	1993	3 (0,0,1,0,0,1,0,1)
4657	Roberts	1986	6 (1,0,1,0,1,1,1,1)
4656	Roberts	1989	6 (1,0,1,0,1,1,1,1)
4806	Roberts	1988	4 (1,0,1,0,0,1,0,1)
3118	Roberts	1988	4 (1,0,1,0,0,1,0,1)
3117	Roberts	1991	2 (0,0,0,0,0,1,0,1)
4319	Roberts	1995	3 (0,0,0,0,0,1,1,1)
2295	Ruben	1997	5 (1,1,0,0,1,1,1,0)
2583	Teele	1990	6 (1,0,1,0,1,1,1,1)

[a] The eight components of study quality score are: (1) study cohort clearly defined; (2) subjects assembled at a uniform time point; (3) pathway of subject entry clearly described; (4) complete follow-up achieved; (5) withdrawals/drop-outs described; (6) objective outcomes used; (7) outcome assessment blinded; and (8) extraneous factors adjusted. '1' indicates presence and '0' indicates absence.

Table 30: Key Characteristics of the Cohort Studies for Question 2

ID#	Author	Year	Cohort	Age of OM History	Age at Outcome Measure	Major Outcome statistic	Outcome Measure	Test	Notes
4675	Owen	1996	Texas	0-3 years	5 years	Correlation	Cognition	Stanford Binet	
4319	Roberts	1995	North Carolina	2 months-3 years	12 years	Correlation	Cognition	WISC-R	
1623	Kaplan	1973	Alaska	0-1 year	10 years	Mean (range)	Cognition	WISC	Stratified by concurrent hearing status
877	Black	1993	Maryland	8-22 months	4-6 years	Mean (SD)	Cognition	McCarthy	
1373	Gravel	1992	New York	0-1 year	4 years	Mean (SD)	Cognition	Stanford-Binet	
4657	Roberts	1986	North Carolina	2 months-3 years	3.5-6 years	Mean (SD)	Cognition	McCarthy	
4656	Roberts	1989	North Carolina	2 months-3 years	8 years	Mean (SD)	Cognition	WISC-R	
4651	Klein	1988	Massachusetts	0-2 years	7 years	Multivariate	Cognition	WISC-R	
2583	Teele	1990	Massachusetts	0-2 years	7 years	Multivariate	Cognition	WISC-R	Adjusted for SES and gender
3118/4806	Roberts	1988	North Carolina	2 months-3 years	2.5-8 years	Correlation	Expressive language	Elicited language play	
3117	Roberts	1991	North Carolina	2 months-3 years	4.5-6 years	Correlation	Expressive language	CELF	Stratified by socioeconomic status
4675	Owen	1996	Texas	0-3 years	5 years	Correlation	Expressive language	Goldman-Fristoe test	
1255	Fischler	1985	Arizona	0-2 years	6-8 years	Mean (SD)	Expressive language	TOLD	
1373	Gravel	1992	New York	0-1 year	4 years	Mean (SD)	Expressive language	SICD-R in months	
2135	Paul	1993	Oregon	0-2 years	4 years	Mean (SD)	Expressive language	MLU	Stratified by normal/late talkers
1219	Feagans	1987	Pennsylvania	0-3 years	5-7 years	Multivariate	Expressive language	MLU, Paraphrase	Reported for total group only
2583	Teele	1990	Massachusetts	0-2 years	7 years	Multivariate	Expressive language	WUG test	Mean adjusted for SES and gender
2295	Ruben	1997	New York	0-1 year	2-9 years, yearly	Percent difference	Expressive language	Unknown	
1277	Freeark	1992	Michigan	0-3 years	3-4 years	Proportion	Expressive language	Verbal Scale Index	Stratified by Parent Verbal Stimulation (PVS)

Abbreviations:

CAVAT= Carrow Elicited Language Inventory
CELF= Clinical Evaluation of Language Functions
MLU= Mean Length of Utterance
PPVT-R= Peabody Picture Vocabulary Test-Revised
SD= Standard Deviation
SICD-R= Sequenced Inventory of Communication Development-Revised
TOLD= Test of Language Development
WISC-R= Wechsler Intelligence Scale for Children-Revised
WRAML= Wide Range Assessment of Memory and Learning
WUG= "WUG" test (Berko-Gleason)

Table 30 (Continued)

ID#	Author	Year	Cohort	Age of OM History	Age at Outcome Measure	Major Outcome statistic	Outcome Measure	Test	Notes
4728	Gravel	1996	New York	0-1 year	9 years	Raw data not reported	Expressive language	WRAML	Only statistical testing was reported
3117	Roberts	1991	North Carolina	2 months-3 years	4.5-6 years	Correlation	Receptive language	CELF	Stratified by socioeconomic status (SES)
3117	Roberts	1991	North Carolina	2 months-3 years	4.5-6 years	Correlation	Receptive language	PPVT-R	Stratified by SES
4675	Owen	1996	Texas	0-3 years	5 years	Correlation	Receptive language	CAVAT	
1255	Fischler	1985	Arizona	0-2 years	6-8 years	Mean (SD)	Receptive language	TOLD	
877	Black	1993	Maryland	8-22 months	4-6 years	Mean (SD)	Receptive language	PPVT-R	
1373	Gravel	1992	New York	0-1 year	4 years	Mean (SD)	Receptive language	SICD-R in months	
2583	Teele	1990	Massachusetts	0-2 years	7 years	Multivariate	Receptive language	WUG test	Adjusted for SES and gender
2295	Ruben	1997	New York	0-1 year	2-9 years, yearly	Percent difference	Receptive language	Unknown	
4728	Gravel	1996	New York	0-1 year	9 years	No raw data	Receptive language	CELF-R	Statistical significance only
3118/4806	Roberts	1988	North Carolina	2 months-3 years	2.5-8 years	Correlation	Expressive speech	Goldman-Fristoe	
2135	Paul	1993	Oregon	0-2 years	4 years	Mean (SD)	Expressive speech	Goldman-Fristoe	Stratified by normal/late talkers
2583	Teele	1990	Massachusetts	0-2 years	7 years	Multivariate	Expressive speech	Goldman-Fristoe	Adjusted for SES and gender
2295	Ruben	1997	New York	0-1 year	2-9 years, yearly	Percent difference	Expressive speech	Unknown	
3118/4806	Roberts	1988	North Carolina	2 months-3 years	2.5-8 years	Correlation	Receptive speech	Articulation tests	
1941	Mody	1999	New York	0-1 year	9 years	Mean (SD)	Receptive speech	Synthetic speech syllables	
1435	Harsten	1993	Sweden	0-3 years	4 years	Proportion	Receptive speech	Linguistic analysis	
4728	Gravel	1996	New York	0-1 year	9 years	No raw data	Receptive speech	Pediatric Speech Intelligibility	Only statistical testing result was reported

Abbreviations:

CAVAT= Carrow Elicited Language Inventory
CELF= Clinical Evaluation of Language Functions
MLU= Mean Length of Utterance
PPVT-R= Peabody Picture Vocabulary Test-Revised
SD= Standard Deviation

SICD-R= Sequenced Inventory of Communication Development-Revised
TOLD= Test of Language Development
WISC-R= Wechsler Intelligence Scale for Children-Revised
WRAML= Wide Range Assessment of Memory and Learning
WUG= "WUG" test (Berko-Gleason)

Table 31: Meta-Analysis for Expressive Language Development

ID Number	Author-Year	Cohort	Age of OM history	Age of outcome measure	Name of Test	Positive OM History			Negative OM History			Standardized Mean Difference (95% CI)
						N	Mean	SD	N	Mean	SD	
1255	Fischler-1985 [a]	Arizona	0-2yr	6-8 yrs	TOLD	33	60.0	20.4	71	64.8	28.8	- 0.18 (-0.59, 0.23)
1373	Gravel-1992	New York	0-1yr	4yrs	SICD-R	8	36.0	5.2	12	39.0	6.2	- 0.49 (-1.40, 0.42)
2135	Paul-1993[a]	Oregon	0-2yr	4yrs	MLU	8	57.8	3.8	13	54.6	10.7	0.35 (-0.54, 1.24)
											Random Effects estimate	-0.14 (-0.49, 0.20)
Test of standardized mean difference equals 0: z=0.82; p=0.413. Test of heterogeneity: Chi-squared=1.77 (degrees of freedom=2); p-value=0.412.												

[a] Retrospective-prospective studies.

Abbreviations:
OM=Otitis media
TOLD=Test of Language Development
SICD-R=Sequenced Inventory of Communication Development-Revised
MLU=Mean Length of Utterance
N=Number of subjects
SD=Standard deviation
CI=Confidence interval

Table 32: Meta-Analysis for Receptive Language Development

ID Number	Author-Year	Cohort	Age of OM history	Age of outcome measure	Name of Test	Positive OM History			Negative OM History			Standardized Mean Difference (95% CI)
						N	Mean	SD	N	Mean	SD	
877	Black-1993	Maryland	8-22 mos	4-6 yrs	PPVT-R	21	83	17	10	72	18	0.62 (-0.15, 1.39)
1255	Fischler [a] - 1985	Arizona	0-2yr	6-8 yrs	TOLD	33	67	28	71	73	32	-0.19 (-0.61, 0.22)
1373	Gravel-1992	New York	0-1yr	4yr	SICD-R	8	36	5	13	38	5	-0.38 (-1.27, 0.51)
2579	Teele-1990	Boston	0-2yr	3yr	PPVT-R	52	101	17	80	96	15	0.31 (-0.04, 0.67)
											Random effects estimate	0.10 (-0.29, 0.49)
Test of standardized mean difference equals 0: z=0.52; p=0.606. Test of heterogeneity: Chi-squared=6.22 (degrees of freedom=3); p=0.102.												

[a] Retrospective-prospective study.

Abbreviations:

PPVT-R=Peabody Picture Vocabulary Test-Revised
TOLD=Test of Language Development
SICD-R=Sequenced Inventory of Communication Development-Revised
N=Number of subjects
SD=Standard deviation
CI=Confidence interval

Table 33: Meta-Analysis for Cognitive Verbal Intelligence

ID Number	Author-Year	Cohort	Age of OM history	Age of outcome measure	Name of Test	Positive OM History			Negative OM History			Standardized Mean Difference (95% CI)
						N	Mean	SD	N	Mean	SD	
877	Black-1993	MD	8-22 mos	4-6 yrs	McCarthy	21	46.7	11.5	10	41.0	10.7	0.49 (-0.27, 1.26)
1373	Gravel-1992	NY	0-1yr	4yrs	Stanford-Binet	9	88.3	15.9	13	84.3	9.4	0.31 (-0.55, 1.17)
4657	Roberts-1986	NC	2mos-3yrs	3.5-6yrs	McCarthy	19	52.0	8.0	19	52.0	9.0	0.00 (-0.64, 0.64)
											Random effects estimate	0.23 (-0.20, 0.65)
Test of standardized mean difference equals 0: z=1.05; p=0.292. Test of heterogeneity: Chi-squared=0.99 (degrees of freedom=2); p=0.609.												

Abbrevations:
OM=Otitis media
N=Number of subjects
SD=Standard deviation
CI=Confidence interval

Table 34: Cohort Studies at Tertiary Screening and Reasons for Exclusion

ID#	Author	Year	Rejection Reason
2221	Rahko	1995	Otitis media not measured at less than 3 years of age
2762	Webster	1989	Case control study
4728	Gravel	1996	No significant difference observed but no hearing data reported
4846	Wright	1984	Hearing data at 3-4 years not abstractable.

Table 35: List of Cohort Studies Included for Question 3

ID#	Author	Year	Cohort
147	Sorri	1995	Birth cohort from Northern Finland
1255	Fischler	1985	Four Indian reservations in Arizona
1373	Gravel	1992	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
1435	Harsten	1993	University Hospital of Lund, Sweden
1623	Kaplan	1973	Eskimo villages in the Yukon and Kuskokwim River Delta areas of Southwestern Alaska
2233	Reed	1967	Eskimo villages in the Yukon and Kuskokwim River Delta areas of Southwestern Alaska
2309	Ryding	1997	University Hospital of Lund, Sweden
2854	Zargi	1992	University of Ljubljana, Ljubljana, Slovenia

Table 36: List of Cohort Studies Excluded from Question 3 (For Reason of Not Reporting Findings Beyond 3 Years of Life)

ID#	Author	Year	Cohort
2264	Roberts	1998	Frank Porter Graham Child Development Center, Chapel Hill, NC
4808	Roberts	1988	Frank Porter Graham Child Development Center, Chapel Hill, NC
667	Abraham	1996	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
2740	Wallace	1988	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
4680	Gravel	2000	LIFE (Longitudinal Infant Follow-up and Evaluation) Program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, Bronx, NY
1288	Friel-Patti	1982	Parkland Memorial Hospital, Dallas, TX
2819	Wright	1988	Pneumococcal vaccine study in Nashville, TN
1677	Knishkowy	1991	PROD (Promotion of Growth and Development) Program, Western Jerusalem
1544	Hutchings	1992	Six general practices in Oxford, England
4838	Vernon-Feagans	1996	Three day-care facilities in central Pennsylvania, PA

Table 37 Definition of Early Life Positive or Negative OM History in Assessing Long-term Hearing Development

ID#	Author Year	Definition of Positive/Negative OM History	Definition of OM	OM Diagnosis Method
147	Sorri 1995	OM history was obtained from all possible sources (health care centers, hospitals, and private surgeries). Only children with clear-cut differences in their history were considered. • RAOM group: had >=4 recurrent episodes until the age of 2 years. • SOM group: had been treated for a long standing (>=3 months) secretory middle ear effusion during the first two years of life. • No OM group: had not experienced an acute otitis media episode until the age of 7 years.	Not specified	Not specified
1255	Fischler 1985	• OM History defined by number of OM episodes by age 2 years and over age 2. • The study defined three groups of children by OM history by number of episodes before and after age 2. • For our assessment, we used groups 1 and 2 as positive OM history, i.e. (>=2 attacks by age 2) and group 3 as negative OM history (<2 attacks by age 2).	Any mention of one of the following: • Acute suppurative OM: history of ear pain or fussiness with or without fever or ear drainage (less than 5 days), and physical evidence of redness with or without immobility, bulging, or a small perforation of the TM. • Serous OM: history of ear fullness, popping, or hearing loss, or an asymptomatic history; and physical evidence of TM retraction and/or immobility, with or without gray or yellow color or bubbles behind the TM • Chronic OM: history of ear drainage and perforation present for more than two weeks; and physical evidence of perforation.	• By medical record review of documented physician's clinical diagnosis.

Table 37 (Continued)

ID#	Author Year	Definition of Positive/Negative OM History	Definition of OM	OM Diagnosis Method
1373	Gravel 1992	OM groups were defined by otoscopic histories during the first year of life. • OM positive: when bilateral OM was detected at 30% or more of the baby's first year visits. • OM negative: when middle ear status was rated as normal in both ears during 80% or more of the first year visits.	Not specified	• Pediatric nurse practitioners completed pneumo-otoscopic examinations during each scheduled well-baby visit
1435	Harsten 1993	OM groups were defined by the number of AOM episodes during the first 3 years of life. • OM positive: developed six or more episodes of AOM during a 12-month period. • OM negative: no AOM episodes and less than six other acute respiratory tract infections.	• AOM was defined as an acute episode of earache in a child with red bulging eardrum(s) or purulent discharge, occasionally febrile and with signs of upper respiratory tract infection.	• By otomicropscopy performed by an otolaryngologist
1623	Kaplan 1973	OM groups were based on age of onset of first episode of otorrhea. • The study defined 3 groups: group 1-onset of first otorrhea episode during first year of life; group 2-onset of first otorrhea episode at 2-10 years, and group 3-no history of OM • For our assessment, we used group 1 as positive OM history and group 3 as negative OM history.	• Used only episodes of OM with otorrhea.	• A research nurse visited the cohort children and obtained information concerning middle ear abnormality and upper respiratory tract illness and reviewed medical records for status between visits.
2233	Reed 1967	Same study population as 1623. This article reported findings at 3-5 years. Article 1623 reported findings at 10 years of age.	Same as 1623	Same as 1623

Table 37 (Continued)

ID#	Author Year	Definition of Positive/Negative OM History	Definition of OM	OM Diagnosis Method
2309	Ryding 1997	OM history groups were defined by the number of recurrent AOM during the first 3 years of life. • rAOM group: children with >=6 episodes of purulent AOM during a 12-month period. • Healthy group: children with no AOM and <6 other RTI episodes during the study period.	• AOM was defined as an acute episode of earache in a child with red bulging eardrum(s) or purulent discharge, occasionally febrile and with signs of upper respiratory tract infection	• AOM was diagnosed by otomicropscopy, performed by an otolaryngologist.
2854	Zargi 1992	OM history based on parental interviews and by review of hospital charts and other medical documentation. • Experimental group: children treated for recurrent acute unilateral or bilateral suppurative OM at 0-2 years of age. • Control group: children who experienced <=1 episode of OM in the first 2 years of life	Not specified	Not specified

Table 38 Study Quality for Studies Included in Evidence Table on Hearing

ID	Author	Year	Study Quality Score[a]
147	Sorri	1995	4 (0,1,0,0,1,1,0,1)
1255	Fischler	1985	5 (1,1,0,0,1,1,1,0)
1373	Gravel	1992	5 (1,1,0,0,1,1,0,1)
1435	Harsten	1993	6 (1,1,1,0,1,1,1,0)
1623	Kaplan	1973	5 (1,1,0,0,1,1,1,0)
2233	Reed	1967	4 (1,1,0,0,0,1,0,1)
2309	Ryding	1997	5 (0,1,1,0,0,1,1,1)
2854	Zargi	1992	2 (1,0,0,0,0,1,0,0)

[a] The eight components of study quality score are: (1) study cohort clearly defined; (2) subjects assembled at a uniform time point; (3) pathway of subject entry clearly described; (4) complete follow-up achieved; (5) withdrawals/drop-outs described; (6) objective outcomes used; (7) outcome assessment blinded; and (8) extraneous factors adjusted. '1' indicates presence and '0' indicates absence.

Table 39: Characteristics and Outcome Measures of Cohort Studies for Question 3

ID Number	Author	Year	Age at OM history	Age at outcome measure	Outcome measured in % hearing loss	Outcome measured in mean pure tone	Other outcome measure
147	Sorri	1995	0-2yrs	7yrs	>20 dB pure tone averages, type not specified	Mean air-conduction (AC) threshold, right/left ear	
1255	Fischler	1985	0-2yrs	6-8yrs	>25 dB at 500Hz; >20 dB at 1000 Hz; >20 dB at 2000 Hz; >25 dB at 4000 Hz; >25 dB at 6000 Hz. Pure tone type not specified.		
1373	Gravel	1992	0-1yrs	4 yrs		Mean pure tone averages obtained at octave frequencies from 500 through 4000 Hz at a minimum, right/left ear	Pediatric Speech Intelligibility sentence (PSI S) to competing messages (CM) ratio
1435	Harsten	1993	0-3yrs	4yrs	>=25 dB tone-audiometry at any frequency, type not specified.		
1435	Harsten	1993	0-3yrs	7yrs	>=25 dB tone-audiometry at any frequency, type not specified.		
1623	Kaplan	1973	0-1yrs	10yrs	>=25 dB air and bone conduction		
2233	Reed	1967	0-2yrs	3-5yrs	>25 dB pure tone air averages		
2309	Ryding	1997	0-3yrs	10yrs		Median level of air conduction hearing, right/left ear	
2854	Zargi	1992	0-2yrs	6-8yrs	>10 dB for air-conduction hearing loss	Sensorineural hearing loss	

Abbreviations:
dB=decibel
PSI=Pediatric Speech Intelligibility
S=Primary sentence
CM=Competing messages

Table 40: Meta-analysis for Effects of Early Life Otitis Media on Long-term Conductive Hearing Loss[a]

Author	Year	Age of OM history	Age at Hearing Testing	OM+ Sample Size	OM- Sample Size	OM+ Percent Hearing Loss	OM- Percent Hearing Loss	Rate Difference in %	95% CI of Rate Difference in %	Risk Ratio	95% CI of Risk Ratio
Sorri	1995	0-2yrs	7yrs	64	35	51.6	20.0	31.6	(13.5, 49.6)	2.6	(1.3, 5.2)
Fischler	1985	0-2yrs	6-8yrs	96	70	9.4	1.4	7.9	(1.5, 14.4)	6.6	(0.8, 50.6)
Harsten	1993	0-3yrs	7yrs	24	56	8.3	5.4	3.0	(- 9.6, 15.5)	1.6	(0.3, 8.7)
Kaplan	1973	0-1yrs	10yrs	162	76	19.8	7.9	11.9	(3.2, 20.5)	2.5	(1.1, 5.7)
Random effects estimates				346	237	21.7	6.4	11.3	(3.3, 19.3)	2.6	(1.6, 4.2)
Test of heterogeneity Chi-square test value						40.3	10.5	7.3		1.1	
Test of heterogeneity Chi-square test p-value						<0.001	0.015	0.064		0.768	

The following analysis excluded article by Sorri.

Author	Year	Age of OM history	Age at Hearing Testing	OM+ Sample Size	OM- Sample Size	OM+ Percent Hearing Loss	OM- Percent Hearing Loss	Rate Difference in %	95% CI of Rate Difference in %	Risk Ratio	95% CI of Risk Ratio
Fischler	1985	0-2yrs	6-8yrs	96	70	9.4	1.4	7.9	(1.5,14.4)	6.6	(0.8,50.6)
Harsten	1993	0-3yrs	7yrs	24	56	8.3	5.4	3.0	(-9.6,15.5)	1.6	(0.3,8.7)
Kaplan	1973	0-1yrs	10yrs	162	76	19.8	7.9	11.9	(3.2,20.5)	2.5	(1.1,5.7)
Random effects estimates				282	202	13.0	4.2	8.4	(3.6,13.2)	2.6	(1.3,5.2)
Test of heterogeneity Chi-square test value						6.8	4.4	1.4		1.1	
Test of heterogeneity Chi-square test p-value						0.034	0.114	0.508		0.566	

[a] Hearing Loss was at >20-25 dB threshold at any frequency with or without treatment measured at 6-10 years of age.
Sorri, Fischler and Harsten did not specify type of pure-tone test used in defining hearing loss. Kaplan used air and bone conduction.
Abbreviations: OM+: positive otitis media history; OM-: negative otitis media history; CI: confidence interval.

Table 41: List of Cohort Studies Included for Question 4

ID Number	Author	Year
759	Avery	1986
766	Babonis	1991
784	Barnett	1998
810	Beery	1975
817	Ben-David	1981
886	Block	1998
888	Bluestone	1973
889	Bluestone	1979
989	Cantekin	1977
990	Cantekin	1980
1238	Fiellau-Nikolajsen	1980
1241	Fiellau-Nikolajsen	1980
1245	Fiellau-Nikolajsen	1983
1250	Finitzo	1992
1280	Freyss	1980
1282	Fria	1980
4879	Fried	1985
4878	Gersdorff	1986
1384	Grimaldi	1976
1397	Haapaniemi	1997
1446	Haughton	1977
3022	Johnson	1980
1600	Jonathan	1989
1632	Karma	1989
1646	Kemaloglu	1999
1650	Kennedy	1982

ID Number	Author	Year
1685	Koivunen	1997
1785	Lovette	1976
1804	Macknin	1987
1817	Mains	1989
1837	Marchant	1986
1936	Mitchell	1990
2012	Nozza	1992
2013	Nozza	1994
2048	Orchik	1978
2049	Orchik	1978
2050	Orchik	1980
2055	Ovesen	1993
2058	Oyiborhoro	1987
2118	Paradise	1976
4790	Paradise	1996
4793	Park	1988
2236	Rees	1992
4804	Renvall	1996
2344	Sassen	1994
2412	Shaw	1978
2545	Szucs	1995
2601	Tom	1994
2607	Toner	1990
2675	van Balen	1994
2713	Vaughan-Jones	1992
2758	Watters	1997

Table 42: Reason for Exclusion of Diagnostic Studies After Full-Length Article Review

ID Number	Author(s)	Year	Rejection Reason
2877	Alho	1998	Data not abstractable
694	Amedee	1995	Diagnostic procedure and gold standard greater than 24 hours apart.
887	Block	1999	Not addressing OME
912	Boswell	1993	No gold standard
968	Buhrer	1985	No gold standard
1015	Chang	1998	Not a diagnostic study
1149	Douniadakis	1993	No gold standard
1167	Duncan	1982	Diagnostic procedure and gold standard greater than 24 hours apart.
1233	Fields	1993	Diagnostic procedure and gold standard greater than 24 hours apart.
1236	Fiellau-Nikolajsen	1979	Diagnostic procedure and gold standard greater than 24 hours apart.
1239	Fiellau-Nikolajsen	1980	Diagnostic procedure and gold standard greater than 24 hours apart.
1281	Fria	1980	Procedure not in scope
4749	Kaleida	1992	Data not abstractable
2014	Nozza	1997	Not all referrals had gold standard
2070	Palmu	1999	Only 42/242 ears had myringotomy
2145	Pellett	1997	No gold standard
2385	Schwartz	1987	No gold standard
2434	Silman	1992	No gold standard
2435	Silman	1994	No gold standard
2438	Silva	1997	Data not abstractable
2442	Silverman	1995	No gold standard
2556	Takahashi	1999	Data not abstractable
2786	Williams	1977	Diagnostic procedure and gold standard greater than 24 hours apart.

Table 43: Number of Articles by Diagnostic Method and Reference Standard

Diagnostic Method	Reference Standard			
	Myringotomy	Tympanocentesis	Validated pneumatic otoscopy	Total
Acoustic reflectometry	6	1	3	10
Audiometry-air and bone conduction thresholds	2			2
Audiometry-air conduction threshold	4	1		5
Binaural micro-tympanoscopy	2			2
Non-pneumatic otoscopy	4			4
Pneumatic otoscopy-examiner validation not specified	3			3
Pneumatic otoscopy-unvalidated examiner	9			9
Pneumatic otoscopy-validated examiner	1			1
Portable tympanometer	8			8
Professional tympanometry	35	6	6	47
Quantitative tympanometry	4			4
Signs/symptoms	2			2
Total	80	8	9	97

Note: A record can be counted more than once within a cell (i.e. controlling for intervention group.)

Table 44: Acoustic Reflectometry (>=5 vs <5 RU) versus Myringotomy

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Fried-1985	44	62	71.0				
	Macknin-1987	84	128	65.6				
	Babonis-1991	68	118	57.6				
	Total	196	308	63.6	64.2	(57.0, 71.5)	3.6	0.168
Specificity	Fried-1985	36	40	90.0				
	Macknin-1987	43	70	61.4				
	Babonis-1991	90	102	88.2				
	Total	169	212	79.7	80.4	(65.0, 95.9)	18.4	<0.001
Prevalence	Fried-1985	62	102	60.8				
	Macknin-1987	128	198	64.6				
	Babonis-1991	118	220	53.6				
	Total	308	520	59.2	59.6	(52.5, 66.7)	5.4	0.067

Table 45: Pneumatic Otoscopy versus Myringotomy

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Paradise-1976	136	138	98.6				
	Cantekin-1977	213	230	92.6				
	Bluestone-197	242	256	94.5				
	Karma-1989	726	753	96.4				
	Mains-1989	102	116	87.9				
	Toner-1990	108	124	87.1				
	Finitzo-1992	107	115	93.0				
	Total	1634	1732	94.3	93.8	(91.4, 96.3)	28.8	<0.001
Specificity	Paradise-1976	56	75	74.7				
	Cantekin-1977	113	140	80.7				
	Bluestone-197	131	169	77.5				
	Karma-1989	277	339	81.7				
	Mains-1989	84	93	90.3				
	Toner-1990	87	98	88.8				
	Finitzo-1992	28	48	58.3				
	Total	776	962	80.7	80.5	(75.1, 86.0)	27.2	<0.001
Prevalence	Paradise-1976	138	213	64.8				
	Cantekin-1977	230	370	62.2				
	Bluestone-197	256	425	60.2				
	Karma-1989	753	1092	69.0				
	Mains-1989	116	209	55.5				
	Toner-1990	124	222	55.9				
	Finitzo-1992	115	163	70.6				
	Total	1732	2694	64.3	62.8	(58.3, 67.2)	30.7	<0.001

Table 46: Portable Tympanometry (Mixed Criteria) versus Myringotomy

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Orchik-1978	21	39	53.8				
	Babonis-1991	92	118	78.0				
	Rees-1992	260	260	100.0				
	Vaughan-Jones-1992	120	135	88.9				
	van Balen-1994	147	156	94.2				
	Koivunen-1997	52	66	78.8				
	Total	692	774	89.4	84.5	(76.0, 93.1)	39.6	<0.001
Specificity	Orchik-1978	35	37	94.6				
	Babonis-1991	84	102	82.4				
	Rees-1992	9	50	18.0				
	Vaughan-Jones-1992	41	65	63.1				
	van Balen-1994	37	77	48.1				
	Koivunen-1997	137	175	78.3				
	Total	343	506	67.8	64.4	(44.3, 84.4)	167.1	<0.001
Prevalence	Orchik-1978	39	76	51.3				
	Babonis-1991	118	220	53.6				
	Rees-1992	260	310	83.9				
	Vaughan-Jones-1992	135	200	67.5				
	van Balen-1994	156	233	67.0				
	Koivunen-1997	66	241	27.4				
	Total	774	1280	60.5	58.5	(40.3, 76.7)	268.9	<0.001

Table 47: Professional Tympanometry (Acoustic Reflex at 500 or 1000 Hz) versus Myringotomy

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Orchik-1978	35	39	89.7				
	Orchik-1980	34	39	87.2				
	Nozza-1992	71	81	87.7				
	Nozza-1994	106	124	85.5				
	Total	246	283	86.9	87.1	(83.2, 91.0)	0.6	0.901
Specificity	Orchik-1978	29	37	78.4				
	Orchik-1980	26	37	70.3				
	Nozza-1992	26	30	86.7				
	Nozza-1994	61	94	64.9				
	Total	142	198	71.7	74.8	(64.6, 85.0)	8.2	0.041
Prevalence	Orchik-1978	39	76	51.3				
	Orchik-1980	39	76	51.3				
	Nozza-1992	81	111	73.0				
	Nozza-1994	124	218	56.9				
	Total	283	481	58.8	58.6	(48.5, 68.6)	14.9	0.002

Table 48: Professional Tympanometry (Static Compensated Acoustic Admittance at 0.1) versus Myringotomy

(A) Including both Nozza articles

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Fiellau-Nikolajsen-1980	9	46	19.6				
	Nozza-1992	25	81	30.9				
	Nozza-1994	37	137	27.0				
	Barnett-1998	95	175	54.3				
	Total	166	439	37.8	33.2	(17.5, 48.9)	38.1	<0.001
Specificity	Fiellau-Nikolajsen-1980	42	42	100.0				
	Nozza-1992	29	30	96.7				
	Nozza-1994	109	112	97.3				
	Barnett-1998	104	124	83.9				
	Total	284	308	92.2	95.0	(88.5, 100)	13.9	0.003
Prevalence	Fiellau-Nikolajsen-1980	46	88	52.3				
	Nozza-1992	81	111	73.0				
	Nozza-1994	137	249	55.0				
	Barnett-1998	175	299	58.5				
	Total	439	747	58.8	59.7	(51.8, 67.7)	14.2	0.003

(B) Excluding Nozza-1992 study

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Fiellau-Nikolajsen-1980	9	46	19.6				
	Nozza-1994	37	137	27.0				
	Barnett-1998	95	175	54.3				
	Total	141	358	39.4	33.9	(12.7, 55.0)	36.9	<0.001
Specificity	Fiellau-Nikolajsen-1980	42	42	100.0				
	Nozza-1994	109	112	97.3				
	Barnett-1998	104	124	83.9				
	Total	255	278	91.7	94.1	(83.9, 100)	13.7	0.001
Prevalence	Fiellau-Nikolajsen-1980	46	88	52.3				
	Nozza-1994	137	249	55.0				
	Barnett-1998	175	299	58.5				
	Total	358	636	56.3	56.3	(52.5, 60.2)	1.4	0.510

Table 49: Professional Tympanometry (Static Compensated Acoustic Admittance at 0.2) versus Myringotomy

(A) Including Both Nozza Studies

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Fiellau-Nikolajsen-1980	21	46	45.7				
	Nozza-1992	74	81	91.4				
	Nozza-1994	63	137	46.0				
	Barnett-1998	110	175	62.9				
	Total	268	439	61.0	61.8	(39.0, 84.7)	93.4	<0.001
Specificity	Fiellau-Nikolajsen-1980	40	42	95.2				
	Nozza-1992	21	30	70.0				
	Nozza-1994	103	112	92.0				
	Barnett-1998	93	124	75.0				
	Total	257	308	83.4	84.5	(74.0. 95.0)	23.0	<0.001
Prevalence	Fiellau-Nikolajsen-1980	46	88	52.3				
	Nozza-1992	81	111	73.0				
	Nozza-1994	137	249	55.0				
	Barnett-1998	175	299	58.5				
	Total	439	747	58.8	59.7	(51.8, 67.7)	14.2	0.003

(B) Excluding Nozza-1992 Study

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Fiellau-Nikolajsen-1980	21	46	45.7				
	Nozza-1994	63	137	46.0				
	Barnett-1998	110	175	62.9				
	Total	194	358	54.2	52.2	(39.5, 64.8)	10.7	0.005
Specificity	Fiellau-Nikolajsen-1980	40	42	95.2				
	Nozza-1994	103	112	92.0				
	Barnett-1998	93	124	75.0				
	Total	236	278	84.9	87.7	(76.8. 98.5)	17.9	<0.001
Prevalence	Fiellau-Nikolajsen-1980	46	88	52.3				
	Nozza-1994	137	249	55.0				
	Barnett-1998	175	299	58.5				
	Total	358	636	56.3	56.3	(52.5, 60.2)	1.4	0.510

Table 50: Professional Tympanometry (Static Compensated Acoustic Admittance at 0.3) versus Myringotomy

(A) Including both Nozza articles

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Orchik-1978	35	39	89.7				
	Fiellau-Nikolajsen-1980	16	46	34.8				
	Nozza-1992	59	81	72.8				
	Nozza-1994	96	137	70.1				
	Total	206	303	68.0	67.4	(49.2, 85.7)	41.6	<0.001
Specificity	Orchik-1978	15	37	40.5				
	Fiellau-Nikolajsen-1980	10	42	23.8				
	Nozza-1992	24	30	80.0				
	Nozza-1994	90	112	80.4				
	Total	139	221	62.9	56.4	(27.5, 85.3)	69.4	<0.001
Prevalence	Orchik-1978	39	76	51.3				
	Fiellau-Nikolajsen-1980	46	88	52.3				
	Nozza-1992	81	111	73.0				
	Nozza-1994	137	249	55.0				
	Total	303	524	57.8	58.2	(48.1, 68.3)	15.9	0.001

(B) Excluding Nozza-1992 study

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Orchik-1978	35	39	89.7				
	Fiellau-Nikolajsen-1980	16	46	34.8				
	Nozza-1994	96	137	70.1				
	Total	147	222	66.2	65.4	(39.1, 91.7)	41.5	<0.001
Specificity	Orchik-1978	15	37	40.5				
	Fiellau-Nikolajsen-1980	10	42	23.8				
	Nozza-1994	90	112	80.4				
	Total	115	191	60.2	48.6	(10.2, 87.0)	64.7	<0.001
Prevalence	Orchik-1978	39	76	51.3				
	Fiellau-Nikolajsen-1980	46	88	52.3				
	Nozza-1994	137	249	55.0				
	Total	222	413	53.8	53.8	(49.0, 58.6)	0.4	0.811

Table 51: Professional Tympanometry (B curve as abnormal) versus Myringotomy

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Orchik-1978	46	84	54.8				
	Shaw-1978	48	49	98.0				
	Johnson-1980	55	74	74.3				
	Ben-David-1981	195	250	78.0				
	Kennedy-1982	44	51	86.3				
	Gersdorff-1986	61	81	75.3				
	Park-1988	182	248	73.4				
	Mitchell-1990	57	65	87.7				
	Toner-1990	107	124	86.3				
	Finitzo-1992	65	72	90.3				
	Vaughan-Jones-1992	91	135	67.4				
	Ovessen-1993	310	342	90.6				
	Sassen-1994	225	279	80.6				
	Tom-1994	100	153	65.4				
	Renvall-1996	87	101	86.1				
	Watters-1997	679	745	91.1				
	Total	2352	2853	82.4	80.9	(76.1, 85.7)	196.6	<0.001
Specificity	Orchik-1978	52	58	89.7				
	Shaw-1978	3	10	30.0				
	Johnson-1980	31	41	75.6				
	Ben-David-1981	39	61	63.9				
	Kennedy-1982	21	24	87.5				
	Gersdorff-1986	27	47	57.4				
	Park-1988	15	38	39.5				
	Mitchell-1990	10	19	52.6				
	Toner-1990	91	98	92.9				
	Finitzo-1992	19	22	86.4				
	Vaughan-Jones-1992	61	65	93.8				
	Ovessen-1993	37	51	72.5				
	Sassen-1994	59	101	58.4				
	Tom-1994	47	60	78.3				
	Renvall-1996	25	26	96.2				
	Watters-1997	166	210	79.0				
	Total	703	931	75.5	74.5	(66.9, 82.0)	147.8	<0.001
Prevalence	Orchik-1978	84	142	59.2				
	Shaw-1978	49	59	83.1				
	Johnson-1980	74	115	64.3				
	Ben-David-1981	250	311	80.4				
	Kennedy-1982	51	75	68.0				
	Gersdorff-1986	81	128	63.3				
	Park-1988	248	286	86.7				
	Mitchell-1990	65	84	77.4				
	Toner-1990	124	222	55.9				
	Finitzo-1992	72	94	76.6				
	Vaughan-Jones-1992	135	200	67.5				
	Ovessen-1993	342	393	87.0				
	Sassen-1994	279	380	73.4				
	Tom-1994	153	213	71.8				
	Renvall-1996	101	127	79.5				
	Watters-1997	745	955	78.0				
	Total	2853	3784	75.4	73.6	(69.1, 78.1)	156.6	<0.001

Table 52: Professional Tympanometry (B or C2 curve as abnormal) versus Myringotomy

					Random Effect Estimate		Test of Heterogeneity	
Measure	Author-Year	X	N	%	%	95% CI	Q	P-Value
Sensitivity	Orchik-1978	31	39	79.5				
	Fiellau-Nikolajsen-1980	42	46	91.3				
	Kennedy-1982	51	51	100.0				
	Vaughan-Jones-1992	120	135	88.9				
	Ovessen-1993	323	342	94.4				
	Sassen-1994	253	279	90.7				
	Total	820	892	91.9	93.8	(91.1, 96.4)	9.44	0.093
Specificity	Orchik-1978	32	37	86.5				
	Fiellau-Nikolajsen-1980	37	42	88.1				
	Kennedy-1982	11	24	45.8				
	Vaughan-Jones-1992	41	65	63.1				
	Ovessen-1993	27	51	52.9				
	Sassen-1994	33	101	32.7				
	Total	181	320	56.6	61.8	(41.5, 82.1)	89.8	<0.001
Prevalence	Orchik-1978	39	76	51.3				
	Fiellau-Nikolajsen-1980	46	88	52.3				
	Kennedy-1982	51	75	68.0				
	Vaughan-Jones-1992	135	200	67.5				
	Ovessen-1993	342	393	87.0				
	Sassen-1994	279	380	73.4				
	Total	892	1212	73.6	67.3	(56.3, 78.2)	90.1	<0.001

Table 53: Summary of Meta Analysis for Diagnostic Comparisons, Excluding Duplicated Studies by Same Author

			Number	Number	Measure	Random Effect Estimate		Test of Heterogeneity	
Measure	**ID**	**Diagnostic Comparison versus Myringotomy**	**Articles**	**Cases**	**%**	**%**	**95% CI**	**Q**	**P-Value**
Sensitivity	**1**	Acoustic reflectometry (>=5 vs <5)	3	308	63.6	64.2	(57.0, 71.5)	3.6	0.168
	2	Pneumatic otoscopy	7	1732	94.3	93.8	(91.4, 96.3)	28.8	<0.001
	3	Portable tympanometry	6	774	89.4	84.5	(76.0, 93.1)	39.6	<0.001
	5	Professional tympanometry (using static compensated acoustic admittance at 0.1)	3	358	39.4	33.9	(12.7, 55.0)	36.9	<0.001
	6	Professional tympanometry (using static compensated acoustic admittance at 0.2)	3	359	54.2	52.2	(39.5, 64.8)	10.7	0.005
	7	Professional tympanometry (using static compensated acoustic admittance at 0.3)	3	222	66.2	65.4	(39.1, 91.7)	41.5	<0.001
	8	Professional tympanometry (using flat or B curve as abnormal)	16	2853	82.4	80.9	(76.1, 85.7)	196.6	<0.001
	9	Professional tympanometry (using flat or B or C2 curve as abnormal)	6	892	91.9	93.8	(91.1, 96.4)	9.4	0.093
Specificity	**1**	Acoustic reflectometry (>=5 vs <5)	3	212	79.7	80.4	(65.0, 95.9)	18.4	<0.001
	2	Pneumatic otoscopy	7	962	80.7	80.5	(75.1, 86.0)	27.2	<0.001
	3	Portable tympanometry	6	506	67.8	64.4	(44.3, 84.4)	167.1	<0.001
	5	Professional tympanometry (using static compensated acoustic admittance at 0.1)	3	278	91.7	94.1	(83.9, 100)	13.7	0.001
	6	Professional tympanometry (using static compensated acoustic admittance at 0.2)	3	278	84.9	87.7	(76.8, 98.5)	17.9	<0.001
	7	Professional tympanometry (using static compensated acoustic admittance at 0.3)	3	191	60.2	48.6	(10.2, 87.0)	64.7	<0.001
	8	Professional tympanometry (using flat or B curve as abnormal)	16	931	75.5	74.5	(66.9, 82.0)	147.8	<0.001
	9	Professional tympanometry (using flat or B or C2 curve as abnormal)	6	320	56.6	61.8	(41.5, 82.1)	89.8	<0.001
Prevalence	**1**	Acoustic reflectometry (>=5 vs <5)	3	520	59.2	59.6	(52.5, 66.7)	5.4	0.067
	2	Pneumatic otoscopy	7	2694	64.3	62.8	(58.3, 67.2)	30.7	<0.001
	3	Portable tympanometry	6	1280	60.5	58.5	(40.3, 76.7)	268.9	<0.001
	5	Professional tympanometry (using static compensated acoustic admittance at 0.1)	3	636	56.3	56.3	(52.5, 60.2)	1.4	0.510
	6	Professional tympanometry (using static compensated acoustic admittance at 0.2)	3	636	56.3	56.3	(52.5, 60.2)	1.4	0.510
	7	Professional tympanometry (using static compensated acoustic admittance at 0.3)	3	413	53.8	53.8	(49.0, 58.6)	0.4	0.811
	8	Professional tympanometry (using flat or B curve as abnormal)	16	3784	75.4	73.6	(69.1, 78.1)	156.6	<0.001
	9	Professional tympanometry (using flat or B or C2 curve as abnormal)	6	1212	73.6	67.3	(56.3, 78.2)	90.1	<0.001

Note: Comparison #4 had only two studies and thus not included in the summary

Table 54: Study Quality and Test Performer of Diagnostic Studies Used in Meta-Analysis

Diagnostic Comparison	Author-Year	Study Quality[a]	Test Performer
1. Acoustic reflectometry (>=5 vs <5 RU) with myringotomy	Fried-1985	1(100000)	Not specified
	Macknin-1987	4(111001)	Pediatrician
	Babonis-1991	5(111011)	One of the authors, specialty not specified
2. Pneumatic otoscopy with myringotomy	Paradise-1976	4(111001)	Pediatrician
	Cantekin-1977	1(100000)	Not specified
	Bluestone-1979	2(100001)	Not specified
	Karma-1989	3(100101)	Otolaryngologist/pediatrician
	Mains-1989	1(100000)	Senior registrar and senior house officer
	Toner-1990	2(100001)	One of the authors, specialty not specified
	Finitzo-1992	4(111001)	Pediatric otolaryngologist
3. Portable tympanometry (mixed criteria) with myringotomy	Orchik-1978	4(111001)	Not specified
	Babonis-1991	5(111011)	One of the authors, specialty not specified
	Rees-1992	1(100000)	Not specified
	Vaughan-Jones-1992	2(100001)	Not specified
	van Balen-1994	4(111010)	General practitioner with special training from ENT department
	Koivunen-1997	3(110001)	Trained nurse
4. Professional tympanometer (using acoustic reflex at 500 or 1000 Hz) with myringotomy	Orchik-1978	4(111001)	Not specified
	Orchik-1980	4(111001)	Not specified
	Nozza-1992	4(111001)	Audiologist/nurses
	Nozza-1994	4(111001)	Clinically certified and licensed audiologist
5. Professional tympanometer (using static compensated acoustic admittance at 0.1) with myringotomy	Fiellau-Nikolajsen-1980	4(111001)	Author
	Nozza-1992	4(111001)	Audiologist/nurses
	Nozza-1994	4(111001)	Clinically certified and licensed audiologist
	Barnett-1998	4(111001)	Research assistant
6. Professional tympanometer (using static compensated acoustic admittance at 0.2) with myringotomy	Fiellau-Nikolajsen-1980	4(111001)	Author
	Nozza-1992	4(111001)	Audiologist/nurses
	Nozza-1994	4(111001)	Clinically certified and licensed audiologist
	Barnett-1998	4(111001)	Research assistant

[a] The first number is the total score of the six components. The six components of study quality score are: (1) appropriate reference standard; (2) test and reference standard assessed independently of each other; (3) blinded reading of results; (4) patient sample included an appropriate spectrum as in clinical practice; (5) reproducibility and interpretation of test results determined; and (6) description of test method sufficient to permit replication. '1' indicates presence and '0' indicates absence of the criterion.]

Table 54: Study Quality and Test Performer of Diagnostic Studies Used in Meta-Analysis (Continued)

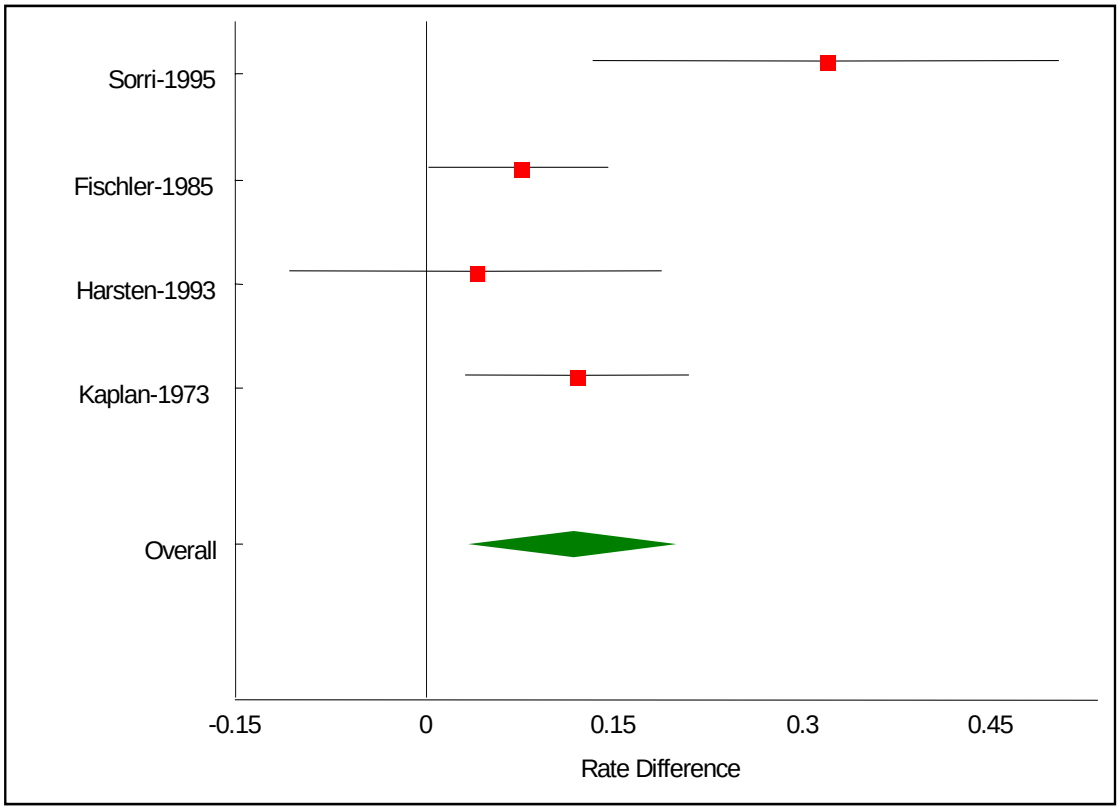
Diagnostic Comparison	Author-Year	Study Quality [a]	Test Performer
7. Professional tympanometer (using static compensated acoustic admittance at 0.3) with myringotomy	Orchik-1978	4(111001)	Not specified
	Fiellau-Nikolajsen-1980	4(111001)	Author
	Nozza-1992	4(111001)	Audiologist/nurses
	Nozza-1994	4(111001)	Clinically certified and licensed audiologist
8. Professional tympanometer (using flat or B curve as abnormal) with myringotomy	Orchik-1978	4(111001)	Not specified
	Shaw-1978	1(100000)	Not specified
	Johnson-1980	2(110000)	Audiologist
	Ben-David-1981	1(100000)	Not specified
	Kennedy-1982	1(100000)	Audiologist
	Gersdorff-1986	1(100000)	Not specified
	Park-1988	1(100000)	Not specified
	Mitchell-1990	1(100000)	Not specified
	Toner-1990	2(100001)	Not specified
	Finitzo-1992	4(111001)	Certified audiologist
	Vaughan-Jones-1992	2(100001)	Not specified
	Ovessen-1993	4(111001)	Otolaryngologist
	Sassen-1994	3(111000)	Not specified
	Tom-1994	4(111001)	Certified audiologists
	Renvall-1996	1(100000)	Not specified
	Watters-1997	2(110000)	Paediatric audiologist
9. Professional tympanometer (using flat or B or C2 curve as abnormal) with myringotomy	Orchik-1978	4(111001)	Not specified
	Fiellau-Nikolajsen-1980	4(111001)	Author
	Kennedy-1982	1(100000)	Audiologist
	Vaughan-Jones-1992	2(100001)	Not specified
	Ovessen-1993	4(111001)	Otolaryngologist
	Sassen-1994	3(111000)	Not specified

[a] The first number is the total score of the six components. The six components of study quality score are: (1) appropriate reference standard; (2) test and reference standard assessed independently of each other; (3) blinded reading of results; (4) patient sample included an appropriate spectrum as in clinical practice; (5) reproducibility and interpretation of test results determined; and (6) description of test method sufficient to permit replication. '1' indicates presence and '0' indicates absence of the criterion.]

Table 55: Country of Origin of Studies Included in Evidence Tables

Country of Origin	All Questions (n=112)	Question 1: Natural History (n=33)	Question 2: Speech and Language (n=20)	Question 3: Hearing (n=7)	Question 4: Diagnostic Methods (n=52)
Belgium	2				2
Canada	1	1			
Denmark	18	14			4
England	9	3			6
Finland	4			1	3
France	1				1
Israel	2	1			1
Italy	1	1			
Korea	1				1
Kuwait	1	1			
Northern Ireland	2				2
Scotland	2	1			1
Sweden	6	3	1	2	
The Netherlands	4	2			2
Turkey	1				1
USA	53	6	19	4	24
not specified	4				4
Non-USA	**55 (51%)**	**27 (82%)**	**1 (5%)**	**3 (43%)**	**24 (50%)**
USA	**53 (49%)**	**6 (18%)**	**19 (95%)**	**4 (57%)**	**24 (50%)**

Figure 1 Shrinkage Plot for Rate Difference in Conductive Hearing Loss at 6-10 Years
Presence of early life OM versus Absence of early life OM



Favors
OM Positive

Favors
OM Negative

Effect size: Difference in proportion of ears with hearing loss
(Positive early life OM – Negative early life OM)

Individual effect size; horizontal line is 95% confidence interval

Pooled effect size; width is 95% confidence interval

Figure 2 Funnel Plot for Rate Difference in Conductive Hearing Loss at 6-10 Years
Presence of early life OM versus Absence of early life OM

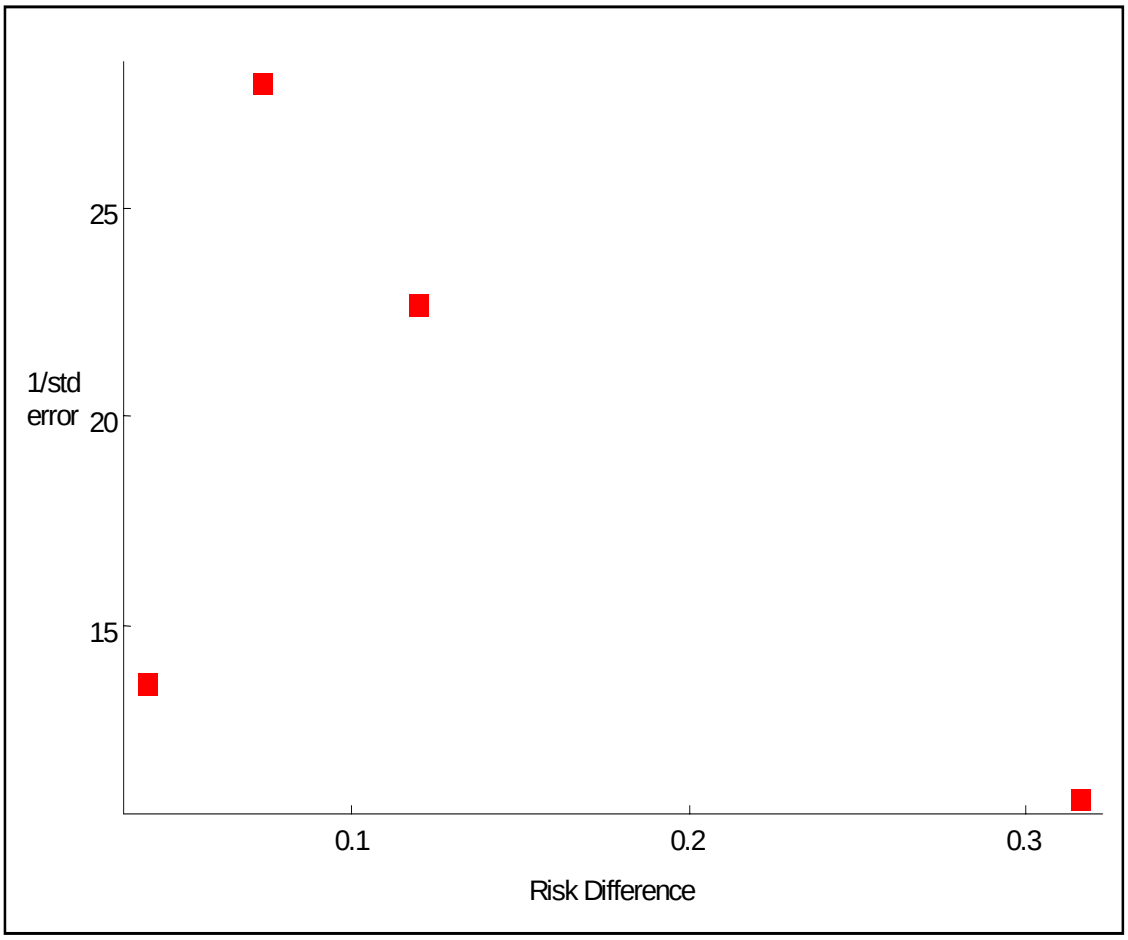
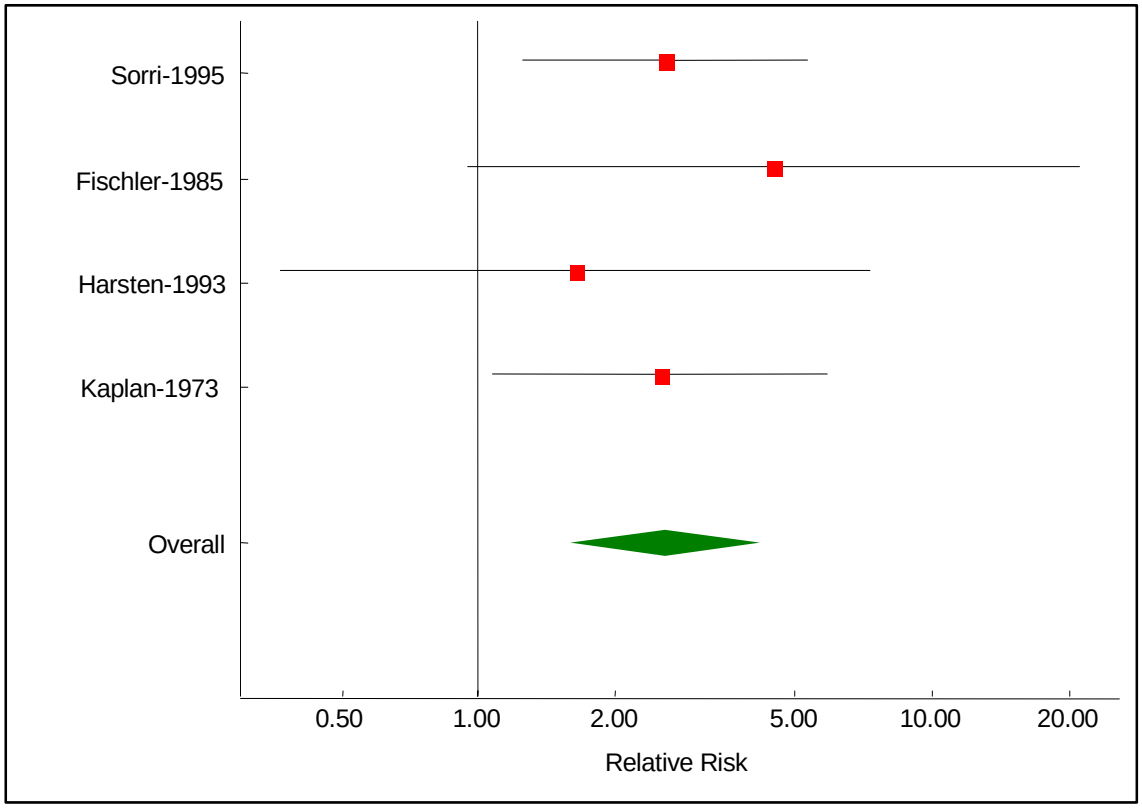


Figure 3 Shrinkage Plot for Relative Risk of Conductive Hearing Loss at Age 6-10 Years Positive Early Life OM versus Negative Early Life OM



Favors OM Positive

Favors OM Negative

Effect size: Difference in proportion of ears with hearing loss (Positive early life OM – Negative early life OM)

Individual effect size; horizontal line is 95% confidence interval

Pooled effect size; width is 95% confidence interval

Figure 4 Funnel Plot for Log Relative Risk of Conductive Hearing Loss at Age 6-10 Years Positive Early Life OM versus Negative Early Life OM

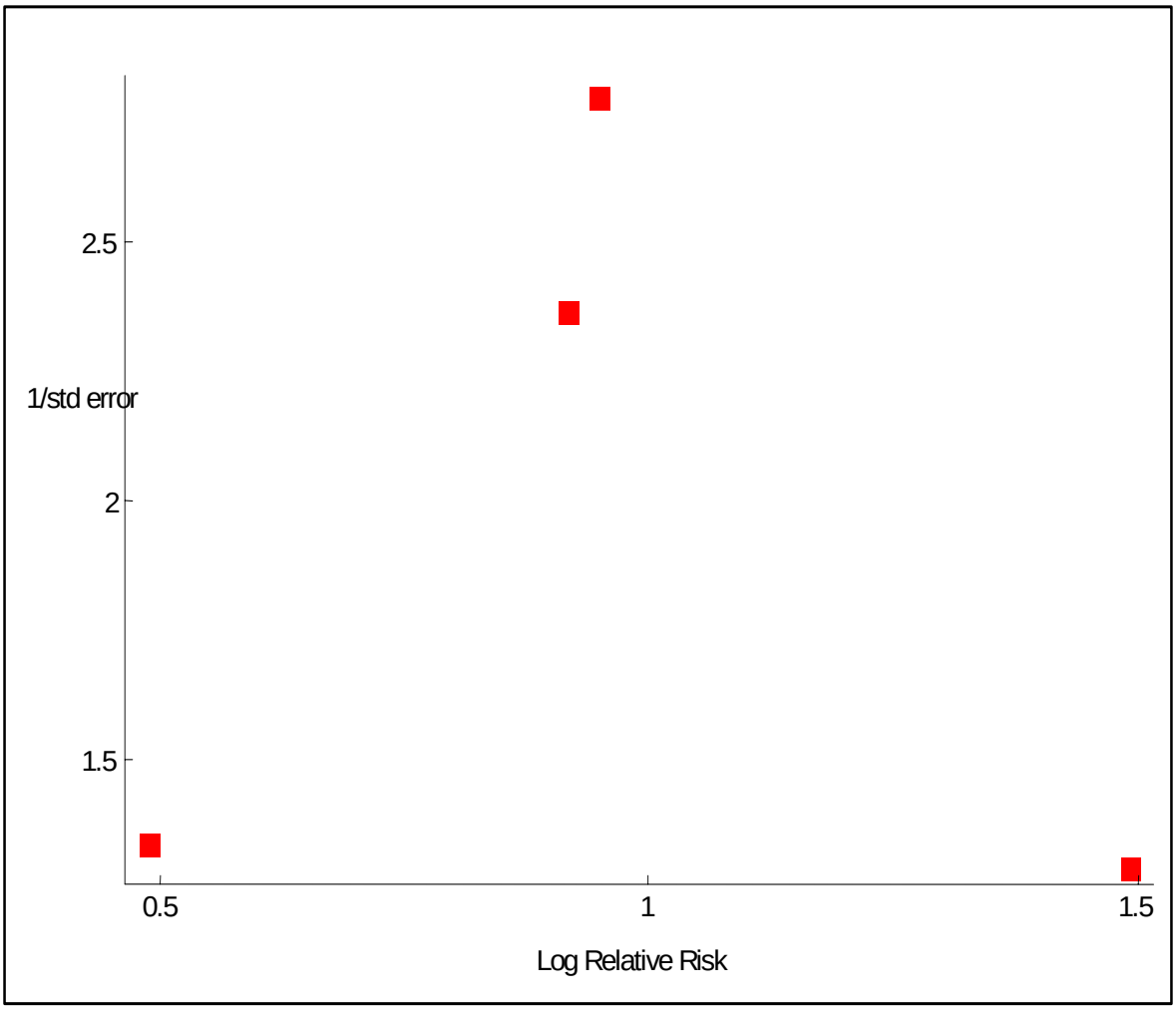


Figure 4 Funnel Plot for Log Relative Risk of Conductive Hearing Loss at Age 6-10 Years Positive Early Life OM versus Negative Early Life OM

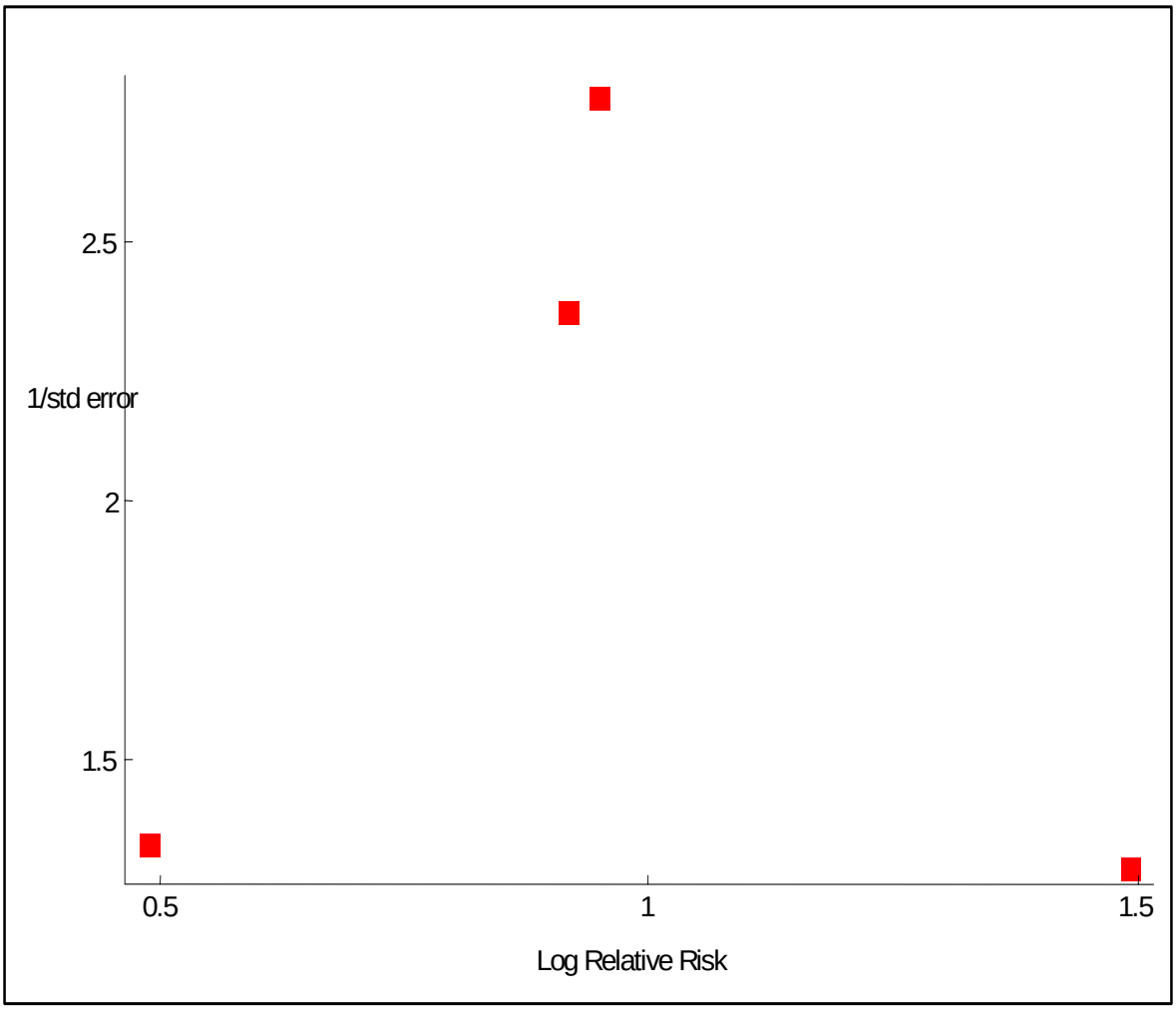
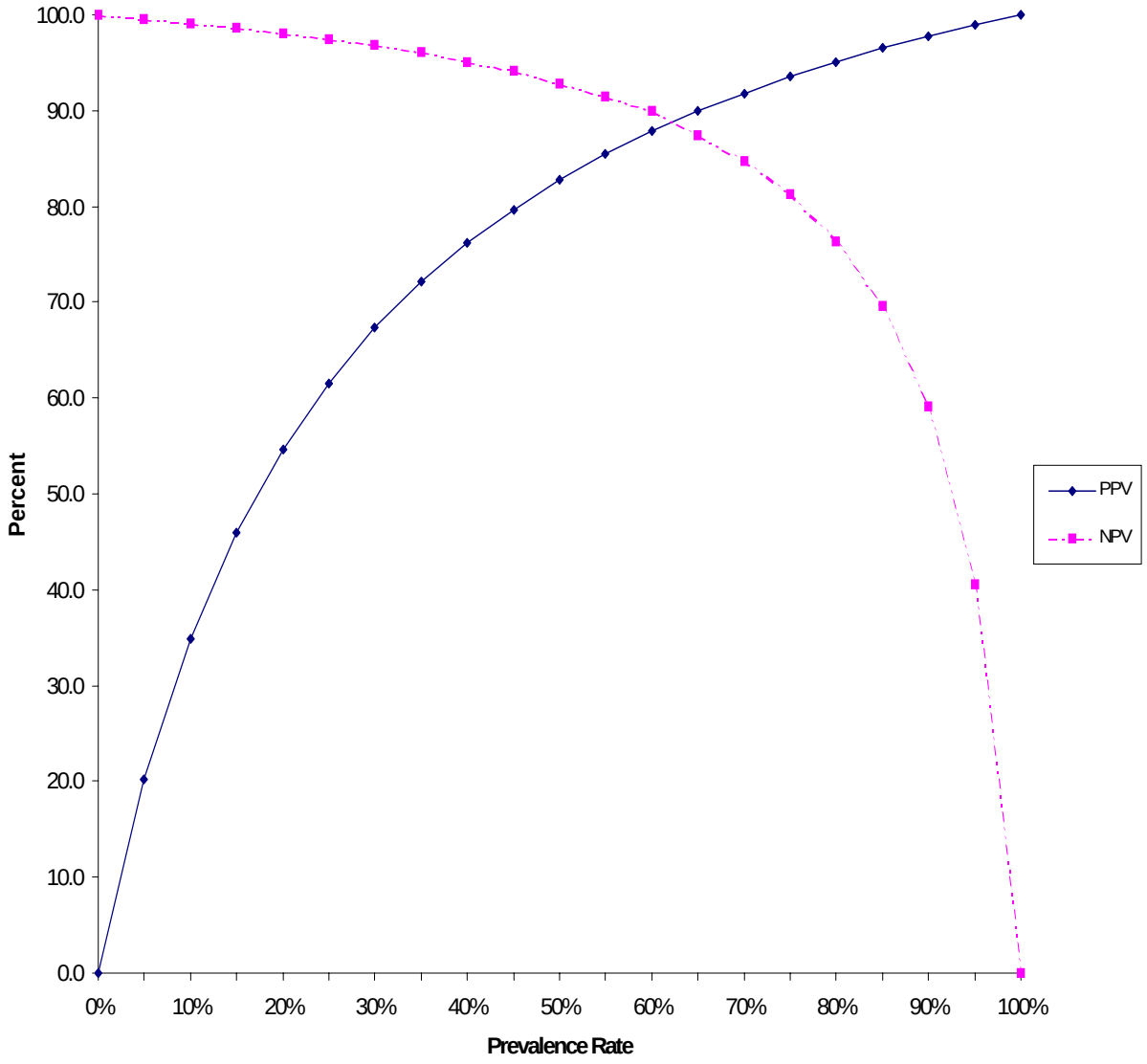


Figure 6 Positive and Negative Predictive Values (PPV and NPV) of Pneumatic Otoscopy by Prevalence of Otitis Media with Effusion Based on Pooled Estimate of Sensitivity at 93.8 % and Specificity at 80.5%



Chapter 4. Conclusions

Despite the recent publication of the 1994 Guidelines on Otitis Media with Effusion in Children (Stool, Berg, Berman, et al., 1994), controversy persists over the management and treatment of otitis media with effusion (OME). The present evidence report systematically reviews the recent literature and provides an update to the 1994 guideline on the diagnosis and the late effects of OME on speech, language, and hearing. In addition, while the 1994 OME guideline (Stool, Berg, Berman et al., 1994) did not formally assess OME resolution rates, the present report reviews the natural history of OME. Although the technical expert panel proposed 20 potential key questions for this report, our time constraints allowed us to address only the four key questions judged by the panel to be most pressing and to have the most significant body of recent literature. We did not address any key questions on treatment or management. Of a total of 449 articles originally identified for this report, only a small percentage qualified for inclusion in the synthesis. Thus, the number of meta-analyses performed was a fraction of those we had hoped would inform this report. Our conclusions should be read with this perspective in mind.

Natural History of OME

Based on the available evidence, this report assessed the natural history of OME in terms of resolution rates. The availability of relevant studies limited our analysis to an assessment of natural history at two periods of followup and only for children 3 years of age and older; we were unable to perform meta-analysis for the under-3 age group. For children older than 3 years of age, we were able to conduct two sets of meta-analyses on the resolution of otitis media with effusion at two followup intervals. These sets were matched by unit of analysis, age group, and OME type and, when possible, diagnostic method.

The first set of meta-analyses assessed resolution at 6 weeks followup. In children older than 3 years of age with OME of unknown duration at the onset of the study, 37.2 or 42.3 percent of ears with OME were free of OME at 6 weeks followup, without regard to their interval OME status, depending on the tympanometric diagnostic criteria for OME resolution. Spontaneous resolution rates were significantly different among the cohorts included in the first set of meta-analyses.

The second set of meta-analyses assessed resolution at 3 months followup. Over a period of 3 months, in studies with cumulative resolution rates, and in children older than 3 years of age with OME of unknown duration, 22.5 or 42.7 percent of ears with OME resolve, depending on the tympanometric diagnostic criteria for OME resolution. A disadvantage of using ears as the unit of analysis is the clinical interpretation of the resolution rate: the distribution of OME resolution using ears cannot be assumed to be the same when children are the unit of analysis. Spontaneous resolution rates were not significantly different among the cohorts included in the second set of meta-analyses.

Apart from age, tympanometric criteria for OME diagnosis, and criteria for resolution on followup, which were reported in the studies, we did not have sufficient information from the studies to consider other influential factors. It was also not possible to know with certainty if the children in the various cohorts presented at identical stages in the course of OME. It is unclear why the non-cumulative, 6-week resolution rate estimate should overlap the cumulative,

3-month resolution rate estimate to such an extent. The non-cumulative resolution rate would be expected to include children whose OME had resolved and then re-accumulated middle-ear effusions (which would make the non-cumulative rate greater than the cumulative rate). Therefore, we view these estimates of OME resolution with great caution.

In the Results section, we also described isolated studies of daily followup, resolution of OME of at least 3 months duration at study onset, and a study that derived OME resolution rate equations. The findings of these studies were too limited to draw any broad conclusions. Similarly, a few studies assessed the effects of influencing factors on OME resolution, but, again, we hesitate to draw any generalizations based on such limited evidence.

Literature on the natural history of OME was difficult to interpret for several reasons. These reasons include its generally poor quality of the research, the lack of control for therapeutic interventions, the inability to distinguish persistent as opposed to recurrent OME due to the length of followup, and the varied criteria for continued followup from examination to examination. We recognize that the meaning of quality summary scores, as measured in this analysis by scoring of documentation in the published articles, can be ambiguous (Jüni, Witschi, Bloch et al., 1999). Nevertheless, whether as a summary score or considering individual quality domains, the quality of the twenty-eight cohort studies on natural history was poor. Twenty-four of twenty-seven of these studies had a quality score of three or less on a scale of one (lowest) to six (highest), even though our preselection process guaranteed all cohort studies a minimum score of one. Half of the studies that attempted to assess the natural history of OME did not control for or document control of interventions, either medical or surgical, that might affect OME outcome during the study period. The majority of these investigations did not stratify findings by intervention status. In addition, the intervals between examinations for OME in these studies varied from one day to 3 years; and, it was impossible to assess whether the "continuing" presence of OME, especially after a long interval, represented persistent or recurrent OME. Furthermore, the criteria for followup varied among studies. Most studies continued followup for the whole study period regardless of OME status at a particular exam, but four cohorts terminated followup of individuals who had type A or normal tympanograms at any exam.

Differing definitions of OME resolution and diagnostic methods also make comparison difficult. As noted in the Results chapter, a difference in the definition of OME resolution between tympanogram type B or C transition to A and tympanogram type B transition to A resulted in a three-fold difference in the estimate of OME resolution in one case. As can be seen from our assessment of the operating characteristics of various OME diagnostic methods in the response to Key Question 4 in the Results, the apparent OME resolution rate could be greatly influenced by the sensitivity and specificity of a test.

A paucity of research studies made it impossible to address several other issues. These included the use of the child or the episode as the unit of analysis. Furthermore, we could not assess OME resolution in younger children or the resolution of OME that was not newly diagnosed and of unknown duration.

The literature was lacking in evidence addressing potential factors that might influence OME resolution. A few studies analyzed the potential influence on OME resolution of factors such as gender, at home care versus daycare, season of onset, side of affected ear, race/et`nicity, or diagnostic instrument. Because of the paucity of such studies, quantitative synthesis was not possible. Such information is also necessary to allow generalizations to specific clinical situations.

Similar to the results of Rosenfeld (1999a), our pooled OME resolution rates could be stratified by age and by tympanometric criteria for OME. Although Rosenfeld did not specify his tympanometric criteria, his analysis found a one-month OME resolution rate of 52 percent (95% CI: 47%, 58%) for OME of unspecified duration in cohort studies with ears as the unit of analysis, regardless of age. This finding is similar to our observation of 42.3 percent resolution (95% CI: 24.1%, 60.6%) for children older than 3 years of age. This finding was not based on cumulative resolution rates but on the disappearance of OME at various points in time in children who had OME at the start of the observation period. Further, Rosenfeld (1999a) found a two-to 3-month OME resolution rate of 63 percent (95% CI: 60%, 63%) for OME of unspecified duration in cohort studies with ears as the unit of analysis, regardless of age. This rate is somewhat higher than our estimate of 42.7 percent (95% CI: 29.3%, 56.1%), which was based on studies that measured cumulative resolution.

We restricted our analysis to prospective cohort studies to provide better assessments of the natural history of OME in the general population of children. Nevertheless, we recognize that information on the natural history of OME may also be gleaned from analysis of the results of randomized, controlled trials of treatment of OME, as Rosenfeld (1999a) subsequently reported. Such estimates would be more applicable than would cohort studies on the general population of children to those who, for whatever reason, are being followed closely for their OME by their health provider. Children being screened for OME in the general population might not necessarily have presented to the health care system for evaluation. As a result, these children might represent a less severe class of OME or at least a different population from those children who were already identified with and being followed for OME and who were deemed potential candidates for therapeutic intervention. Rosenfeld (1999a) reported OME resolution rates of 12 percent (95% CI: 8%, 16%), 23 percent (95% CI: 21%, 26%), and 24 percent (95% CI: 17%, 32%), at two weeks, one month, and one to 3 months followup, respectively, in children with OME that had lasted weeks or months and were randomized to the placebo or no-drug arms of randomized controlled trials. In addition, he assessed OME resolution rates of children with OME that lasted 3 months or longer who were surgical candidates for tympanostomy tubes. He found the following OME resolution rates at 6 months, 1 year, 2 years, 3 years, and 4 years for children assigned to the no intervention groups: 27 percent (95% CI: 20%, 35%), 32 percent (95% CI: 24%, 40%), 31 percent (95% CI: 25%, 37%), 49 percent (95% CI: 37%, 62%), and 59 percent (95% CI: 46%, 72%). A strength of these studies is that the diagnosis of OME tended to rely on pneumatic otoscopy as well as tympanometry, rather than on tympanometry alone. A major difficulty of these studies, as with the cohort studies, is the inability to account for interventions received from other providers and beyond the control of the investigators. In fact, one of these investigators mentioned that “Antibiotic treatment was not controlled for in this study” (Maw and Bawden, 1994). Another difficulty, as with the cohort studies, is the inability to distinguish persistent from recurrent OME, due to the long intervals between followup assessments.

Finally, the publication of multiple articles based on the study of a single cohort raises a complex issue. If multiple studies from one cohort were included once in a meta-analysis, the results would be unbiased. However, if findings from a single cohort were included more than once in a meta-analysis, then bias would exist.

Effects of Early-life OM on Long-term Speech and Language Development

Our intent was to examine the influence of long and short-duration OME that results in hearing loss on delays in speech and language development and the ability of other factors to modify the effect of OME. However, we found few studies that addressed these questions. Only 20 of 112 studies (19%) (on 12 cohorts) met our criteria for consideration for this question. These studies further suffered from lack of uniformity with respect to risk factors studied, type of outcome measured, method of measurement, unit of measurement, and age at outcome determination.

The generalizability of our findings also is questionable, since nine of the twelve cohorts primarily included children from specific ethnic/racial groups or particular socioeconomic groups. Five of the six studies included in the three meta-analyses were based on such cohorts. However, we excluded studies on children with specific medical conditions such as craniofacial defects, primary mucosal disorders, immunodeficiencies, and genetic disorders. The available literature did not address the effect of pre-existing speech, language, or other developmental disorders, and only one study focussed on the effect of bilateral persistent middle-ear effusion.

Based on our criteria, only nine of the twenty studies (45 percent) were of acceptable quality. These studies scored five points or more on an eight-point scale.

Our analysis was limited to assessing whether early-onset OME resulted in delays in expressive and receptive language development and cognitive verbal intelligence. For children older than 3 years of age who had a positive history of otitis media during the first 3 years of life, our meta-analyses on long-term expressive language (three studies), receptive language (four studies), and cognitive verbal intelligence (three studies) showed no effect of early otitis media. The 1994 OME guideline (Stool, Berg, Berman et al., 1994) had determined that meta-analysis of the 14 “adequate” studies was not possible, because of the wide variety of measurement tools used for outcome assessment as well as a lack of standardization of data reporting. We were able to identify articles that had appeared since the release of the 1994 OME guidelines.

Nevertheless, our findings on the possible effects of early-life OM on speech and language development are in general agreement with the conclusion of the 1994 OME guideline (Stool, Berg, Berman et al., 1994). They concluded that " rigorous, methodologically sound research does not adequately support or refute the theory that untreated otitis media with effusion results in speech/language delays or deficits." Our findings were also in agreement with those noted in the 1994 OME guideline that "Conflicting findings among studies can be accounted for in several ways: limitations in the research designs, lack of uniformity of test instrument selection, lack of definition of hearing status, and interactions between otitis media with effusion and other risk factors," as well as differences in populations studied. Thus, we caution clinicians not to generalize these findings to children with the underlying chronic medical conditions that were excluded from this study or to those with pre-existing developmental disorders. In addition, generalization to children with persistent bilateral OM may not be valid, since only one of the studies specifically assessed bilateral as opposed to unilateral or bilateral otitis media.

Several ongoing, prospective studies are assessing the effect of early otitis media on long-term speech and language development of children older than 3 years. However, the results of these studies have not yet been published (Paradise, Dollaghan, Campbell et al., 2000; Feldman, Dollaghan, Campbell et al., 1999; Paradise, Rockette, Colborn et al., 1997; Roberts, Burchinal,

Zeisel et al., 1998; Roberts, Burchinal, Jackson et al., 2000). These studies will provide further data on which to base an assessment of the effect of OM on speech and language development.

Effects of Early-life OM on Long-term Hearing

Although the immediate effect of OM on the conductive aspects of hearing is well recognized, the long-term effect of early OM on sensorineural hearing has not been well established (Madell, 1999). Our findings indicate that early OM may have an effect on long-term hearing.

Of the eight cohort studies we analyzed, four reported percent of hearing loss at 6 to 10 years of age. Children with early-life OM have a 2.6 times higher risk of hearing loss at 6 to 10 years of age than do children with no early history of OM with an estimated rate difference of 11 percent. The rate difference and the risk ratio were not significantly different among the cohorts. If early OM does indeed lead to long-term hearing deficits, the clinical implications are significant. Depending on the degree of hearing loss, the hypothesized effect of OM on speech and language development as mediated through hearing loss may be of greater duration than would be explained by the intermittent and transient episodes of conductive hearing loss associated with OM. In addition, the question of the effect of OME treatment on long-term hearing gains greater relevance.

The literature available for assessment of the long-term effects of early-life OM on hearing is both limited and of poor quality. Only four percent (eight of the 186) of the studies that addressed OM and hearing qualified for inclusion in our analysis, and only half of these were of acceptable quality. The evidence on the effects of early OM on long-term hearing also suffered from the same methodological issues as the evidence on the effects of early-life OM on long-term speech and language development. Because of the limited nature of this evidence, and because the rate of intervention depends greatly on the threshold hearing level adopted, the findings of our analysis should be used with caution.

Diagnostic Methods for OME

Previous assessments of diagnostic techniques for OME have recommended several different techniques. The 1994 OME Guidelines recommended pneumatic otoscopy (Stool, Berg, Berman, et al., 1994).

Using 52 diagnostic studies, we were able to evaluate the diagnostic accuracy of the following eight methods: acoustic reflectometry at ≤5 or >5 RU; pneumatic otoscopy; portable tympanometry; professional tympanometry using static compensated acoustic admittance at 0.1, 0.2, and 0.3; professional tympanometry using B curve as abnormal; and professional tympanometry using B or C2 curves as abnormal. All comparisons used myringotomy as the reference standard.

Among the eight diagnostic methods, the receiver-operator characteristic points (plotting sensitivity against [1 minus specificity]) for pneumatic otoscopy were closest to the optimal point of 100% sensitivity and 100% specificity. The pooled sensitivity for pneumatic otoscopy was 94 percent (95% CI: 91%, 96%); the pooled specificity was 80 percent (95% CI: 75%, 86%); the positive predictive value was 89 percent (95% CI: 87%, 92%); the negative predictive value was 89 percent (95% CI: 86%, 93%), and the accuracy was 89 percent (95% CI: 87%,

91%). These findings were based on 2,694 children from seven studies that reported a pooled prevalence rate for OME of 63 percent (95% CI: 58%, 67%). The prevalence rates among the studies ranged from 56 percent to 71 percent, which was a significant variation ($p<0.001$). The diagnostic test with the highest specificity was professional tympanometry (using static compensated acoustic admittance at 0.1) at 95.0 percent (95% CI: 88.5, 100).

Our findings are in general agreement with the recommendations and findings of the 1994 OME guideline (Stool, Berg, Berman et al., 1994). The 1994 OME guideline stated that "The diagnostic evaluation of suspected otitis media with effusion should include pneumatic otoscopy. Otoscopy alone (without the use of the pneumatic otoscope to test tympanic membrane mobility) is not recommended." This recommendation was deemed a strong one based on limited scientific evidence and strong Panel consensus. Based on limited scientific evidence and expert opinion, the 1994 OME guideline also allowed the following option: "Tympanometry may be used as a confirmatory test for otitis media with effusion." The OME Guidelines did not include quantitative syntheses of the evidence.

Our analysis considered several references not cited by the 1994 OME guideline; some were accepted, and some were rejected. The results of our meta-analyses confirm that pneumatic otoscopy had the best operating characteristics among the nine alternatives examined. Our findings also confirm that certain, but not all, categories of tympanometry also perform well in identifying middle-ear effusion in OME as well as in distinguishing it from other entities. While the 1994 OME guideline did not make a recommendation regarding acoustic reflectometry, our findings suggest that acoustic reflectometry does not perform as well as pneumatic otoscopy and certain types of tympanometry. The use of the spectral gradient angle, as the unit of measurement, may improve the sensitivity of acoustic reflectometry compared to the use of reflectivity, but this observation is based on a single study that found a sensitivity of 95.4 percent using a threshold of 95 degrees (Barnett, 1998) (compared to the pooled sensitivity of 64.2 percent for a threshold reflectivity of 5 found in this study. However, the specificity was 31.5 percent when a spectral gradient angle of 95 degrees was used as the threshold, compared to the pooled specificity of 80.4 percent with a threshold reflectivity of 5 in this study. Unlike the 1994 OME guideline, which commented on the combination of tympanometry and pneumatic otoscopy, we did not assess combinational diagnostic methods or algorithms.

The finding that pneumatic otoscopy can do as well as or better than tympanometry and acoustic reflectometry has significant practical implications. For the typical clinician, pneumatic otoscopy should be easier to employ than other diagnostic methods. The important question may be what degree of training will be needed for the clinician to be as effective with pneumatic otoscopy as were the examiners in the studies reviewed in this report. Also, while we did not do a cost-effectiveness analysis, the cost of pneumatic otoscopy, in terms of direct and indirect costs, would appear to be less than that for tympanometry or acoustic reflectometry.

Because of inadequate evidence, we could not conduct evaluations of clinical signs and/or symptoms, air and/or bone threshold audiometry, binocular micro-tympanoscopy, and non-pneumatic otoscopy in the diagnosis of OME. In addition, diagnostic methods that use algorithms or aggregated scorings are important but were not included in the scope of this evidence assessment. The sources of variation of such combinational methods are difficult to detect in published articles. In addition, we must emphasize that we assessed the diagnosis of OME middle-ear effusion at single points in time rather than the diagnosis of persistent or recurrent OME over time. The meta-analyses of diagnostic tools raised several methodological concerns. One concern centered on pooling data from studies of diverse populations. The

differences in OME prevalence among the studies point to one aspect of this diversity. In addition, we were concerned about the different models of instruments within the broad categories of diagnostic methods. We assume instrumentation has improved or at least changed over time in the areas of tympanometry and acoustic reflectometry as well as in the other diagnostic methods that could not be assessed in depth.

In addition, as alluded to above, a definitive assessment of diagnostic methods for the diagnosis of middle-ear effusion in OME would require an assessment of the cost-effectiveness of the different diagnostic procedures. Such an assessment could not be incorporated into our analysis due to a limited time frame. Cost-effectiveness analysis would take into account the specific impact of test results, including false positives and false negatives, which will depend on examiner proficiency and the patient preferences for specific outcomes. Cost-effectiveness analysis would also establish optimal operating points or thresholds for diagnostic methods measured on both ordinal and continuous scales (Sox, Stern, Owen et al, 1989).

Study Limitations

Several limitations of the study applied to the evidence analysis as a whole. These included the issues of the selection of key questions, publication language as an exclusion factor, study quality, and the analysis of influencing factors. This section addresses these issues. Limitations that pertained to individual questions are discussed in the Results section.

Question Selection

Several peer reviewers commented on the absence of OME treatment as a topic of inquiry of this study. The key questions were selected using a standard consensus methodology, as described in the Methods section. The technical experts were asked to consider various factors when selecting questions. One of these factors was the importance of the question (as assessed by the potential impact on OME outcomes and on development of future OME guidelines). Another consideration was feasibility (as assessed by the ability to complete the study in 6 months and the availability of sufficient information or new information, if a systematic review had already been done in the past). Therefore, treatment questions were not arbitrarily included or avoided; instead, adherence to our key question selection method led to the questions that were included. Whether the methodology for selecting key questions needs to be reconsidered in future systematic reviews is a valid question. In addition, whether the composition of a particular technical expert panel might influence the selection of key questions is also a valid question for future systematic reviews.

Study Language

For this analysis, we did not include literature published in other languages, although we did consider studies conducted in non-English-speaking nations but published in English-language journals. This decision was based on our previous experience of limited yield from non-English language publications for our evidence assessment of the management of acute otitis media (Takata, Chan, Shekelle et al., in press; Chan, Takata, Shekelle et al., in press). In that assessment, we reviewed a total of 97 articles published in non-English-language journals and

found two eligible for inclusion. However, these studies were also reported in English language publications. In **Table 55**, we tabulated the country of origin of the 112 studies included in the evidence tables. Of the 108 studies that specified country of origin, 55 (51%) were from (15) countries other than the United States. The percentage of studies from countries other than the United States was 82 percent (27 of the 33) for the natural history question, 50 percent (24 of 48) for the diagnostic methods question, 43 percent (three of seven) for the hearing question, and 5 percent (1 of 19) for the speech and language development question. We also observed that among our peer reviewers, four of whom were European (including two from non-English-speaking countries), none mentioned any specific studies from the non-English-language literature that they believed should have been included in the analysis, based on our study criteria.

Study Quality

For two reasons, no prospective cohort studies were excluded based on study quality. First, we decided that if sufficient numbers of studies were available, we would study the variability of findings by either stratified or sensitivity analyses. Secondly, we recognized the potential problems with summary quality scores (Jüni, Witschi, Bloch et al., 1999). For example, we recognized that a study may have adhered to a high quality of study design but not have documented that design in the article. The abstracted data may be viewed along with our quality score, based on the design described in the article, in the evidence tables.

Influencing Factors

As we mentioned in the Results section, we were unable to conduct in-depth analysis by influencing factors for any of our key questions, due to the constraints of the available evidence. In particular, our assessment of the effect of early-life otitis media on long-term speech, language, and hearing was limited to an assessment of the effect of OME duration. The available evidence was not sufficient to allow us to address the second part of the question, namely, the influence of other risk factors, using standard analytic techniques within the resources and timeframe of this evidence analysis. The technical expert panel had listed many demographic (including SES), environmental, and clinical factors that might either act independently or interact with OME to affect speech and language. To address this question, a meta-regression approach would be required to identify the risk factors that contribute significantly to speech, language, and hearing delays in the context of otitis media and could include both comparative and single cohort studies. However, many issues must be addressed in order to set up data for meta-regression analysis appropriately. Such issues would require a great deal more input from technical experts than was possible with the resources available in this study.

Chapter 5. Future Research

General Issues

The need for future research on otitis media with effusion (OME) and otitis media (OM) can be substantiated only by the demonstration of a negative effect of OM on important outcomes such as speech, language, and hearing. The available data do not provide such evidence in the case of speech and language and, suggest, at most, a possible effect of early-life OM on long-term hearing, based on evidence that may not be generalizable. Thus, future research must still establish the effect of early-life OM on speech, language, and hearing. Such future research will benefit by addressing the following general issues, which affect study quality and outcomes assessment. The definitions of OME and OM and of relevant interventions, influencing factors, and outcomes should be standardized.

A common, testable framework, with flexibility for competing hypotheses, that links predisposing factors to OME and OM and then OME or OM to outcomes such as long-term speech and language development and hearing, should be adopted. This framework should include hypotheses on the role of child characteristics, environmental and social influences such as socioeconomic status, and medical factors such as interventions, on outcomes.

Agreement on appropriate follow-up intervals to provide valid estimates of duration or frequency of OME and OM is needed to help in comparing results from different studies.

Additional areas where future research should focus include potential gaps in practice and newer outcome measures. Such outcome measures include general health status and quality issues such as satisfaction with treatment.

Natural History

Future research on the natural history of OME must focus on improvement of study quality and establishing the effect of OME on long-term outcomes such as speech, language, and hearing. In particular, control of therapeutic intervention during the study and the distinction between OME persistence and recurrence need to be addressed. Considering the difficulties of conducting a natural history study on OME, a less restrictive definition of nonintervention might be considered. Even with a less restrictive definition of nonintervention, studies should consider presenting data that has been stratified by the level of intervention each child receives during the study period. For example, if a child is allowed to have antibiotics, the exact circumstances when antibiotics may be given should be determined a priori, the number of episodes of such antibiotic administration should be noted, and the outcome measures reported should be stratified by intervention level. However, with a less restrictive definition of nonintervention, the researcher runs the risk of conducting a study with little meaning or applicability to natural history.

Researchers and clinicians should agree upon standard procedures for follow-up, including intervals of follow-up, so that resolution rates are indeed comparable. As the study by Moller and Tos (1990) demonstrated, even daily exams did not necessarily lead to a greater distinction between persistence of OME and recurrence; thus, we do not expect this to be an easy issue to resolve. The issue of assessment of OME duration or recurrence is as important as the issue of diagnosis of OME at a single point in time.

Researchers need to agree upon a definition of OME resolution and the diagnostic methods with which to make that determination. The latter will require further research into the operating characteristics of OME diagnostic methods, as we have reported in this evidence report. Further enhancements in diagnostic methods must also be achieved. Although perhaps more relevant to studies on outcomes other than resolution, more research is needed that views the child as the unit of analysis, since the actual outcomes of concern, such as speech, language, and hearing, are functional requirements of a child, not an ear.

Further research on the role of influencing factors, such as socioeconomic status, on the natural history of OME may help the clinician in a particular setting make a better decision when assessing a child with particular characteristics. Among the influencing factors the technical expert panel thought were potentially important, the only factors addressed were age, gender, daycare setting, season, and racial/ethnic origin, and those were addressed by only a few studies. Study findings will be useful when they are generalizable, either because of their similarity to the population served by a particular clinician or because they address children with characteristics similar to a specific child the clinician is assessing. However, the issue of OME resolution is relevant only if OME has an impact on outcomes of relevance, such as long-term speech, language, or hearing.

Effects of Early-life OM on Long-term Speech and Language Development

For evaluation of long-term effects of early-life otitis media on speech and language development, a coordinated uniform approach that uses a rational conceptual framework is recommended. Such an approach should address the risk factors, such as socioeconomic status, interventions, and outcome measures in an integrated fashion. Conceptual frameworks include the Global Language Model, the Interactive Language and Attention Model, and a transactional model (Vernon-Feagans, 1999; Roberts and Wallace, 1997). The Global Language Model hypothesizes that mild to moderate hearing loss that results from otitis media is the actual causal factor that leads to speech and language deficits. The Interactive Language and Attention Model also hypothesizes an important role for hearing loss. However, this model also distinguishes between early and later developmental effects and the timing of hearing loss in the life of the child and places greater emphasis on risk and protective factors that may interact with hearing loss as well as directly affecting speech and language development. The transactional model ascribes an important role to differences in parent or caregiver response to children with and without chronic otitis media on long-term effects.

Generalizability of study findings will be enhanced in future research if details of risk or influencing factors and interventions are well planned and documented. Included in this assessment of risk factors should be the issue of hearing loss (both conductive and sensorineural) associated with otitis media as a possible cause of long-term speech and language deficits.

For future systematic reviews, we propose the consideration of an “individual-level-data meta-analysis” method (Mathew and Nordstrom, 1999; Stewart and Clarke, 1995; Stewart and Parmar, 1993) to study the long-term effects on outcomes such as speech, language, or hearing, with many of the suggestions for improvement of study quality noted above. This approach would call for the collaboration of investigators from various institutions who have been following cohorts of children prospectively to contribute data on individual members of their cohorts. Eligible cohorts are identified based on *a priori* criteria. Risk factors, interventions,

and outcomes of interest are also defined *a priori*. The unique feature of "individual-level-data meta-analysis" is the ability it confers to retrieve a uniform set of data directly on risk factor, intervention, and outcome data, case by case. This case-specific data set could then be analyzed using meta-regression or other multivariate techniques. A meta-analysis of updated individual patient data has been found to provide the least biased and most reliable means of addressing questions that have not been satisfactorily resolved by individual studies (Stewart and Parmar, 1993). However, the quality of data and the ability for cohort investigators to collect and share relevant data are important factors in the success of this approach (Steward and Clarke, 1995).

In all aspects of analysis, definitions, classifications, and types and units of measure should be developed by a team of experts prior to the start of a study. A consensus on the definitions and classification of otitis media and on relevant outcome measures will allow for comparisons among cohorts. We realize the lack of knowledge as to what specific aspects of speech and language development might be affected by early otitis media. Nevertheless, we would encourage experts to develop a uniform panel of tests that would measure the broad array of possible aspects of speech and language development hypothesized to be affected and could be consistently applied in research studies by all investigators. Literature on findings should report univariate as well as multivariate analysis findings to allow pooling of data. Many studies reported correlation coefficients or regression coefficients, which are difficult to interpret and to use in quantitative synthesis.

Several prospective studies on the effect of early otitis media on long-term speech and language development are ongoing (Paradise, Dollaghan, Campbell et al., 2000; Feldman, Dollaghan, Campbell et al., 1999; Paradise, Rockette, Colborn et al., 1997; Roberts, Burchinal, Zeisel et al., 1998; Roberts, Burchinal, Jackson et al., 2000). Whether these studies answer more definitively the questions regarding the effect of otitis media on long-term speech and language and delineate areas, apart from the general research issues noted above, for further prospective studies on speech and language will be better assessed when the results of these studies are reported in the peer-reviewed literature.

Effects of Early-life OM on Long-term Hearing

Future research should attempt to confirm whether early-life otitis media leads to more permanent hearing loss than intermittent and transient conductive hearing loss. The importance of hearing loss, whether intermittent and transient or permanent and long-standing, associated with early-life otitis media, should be addressed as noted in the section above on the effect of early otitis media on speech and language development. Similar methodological recommendations, including the "individual-level-data meta-analysis" approach, apply to research on long-term hearing effects and speech and language effects of early otitis media. If OM does affect long-term hearing, the effect of OM treatment on long-term hearing and its cost-effectiveness are of great importance and must be addressed in future prospective studies.

Diagnostic Methods for OME

Future research on the diagnosis of OME will need to start with the definition of OME. The difficulty in reaching a consensus on the definition of OME was seen in our discussion of this issue with our technical expert panel. The technical experts agreed that OME was defined as "fluid in the middle ear without signs or symptoms of ear infection," as proposed by the 1994

OME guideline (Stool, Berg, Berman et al., 1994). However, they could not agree on which signs or symptoms should be absent, i.e. what signs or symptoms differentiated OME from acute otitis media. Without such agreement, we believe little progress can be made in improving the diagnosis of OME.

Limiting the assessment of OME diagnostic methods to those that address middle-ear effusion specifically, as we did, will require further expert consensus on important conceptual issues. One issue that was brought to our attention by one of our technical experts and that was discussed in depth by our technical expert panel was whether diagnosis of middle-ear effusion in the child with OME was different than in the child with acute otitis media. For example, since the child with acute otitis media is in discomfort, whatever the symptoms that are ascribed to acute otitis media, that child will be more difficult to examine for the presence of middle-ear effusion than a child with OME, who by definition is asymptomatic. After much discussion, our panel decided that the diagnosis of middle-ear effusion was different in the context of these two clinical conditions; however, we are aware that other experts may not agree with this opinion. Future systematic reviews will require studies of much higher quality than are currently available. In addition, future studies must provide details on the characteristics of the children studied and the study setting so that the generalizability of the findings can be assessed. Studies confined to children with known middle-ear effusions in tertiary care settings may be easier to conduct, but the clinician in general practice is faced with children whose middle-ear status is unknown at the time of presentation. Future research must provide information that is applicable to the child with unknown middle-ear status in the primary care setting.

Pneumatic otoscopy might appear to be less costly and more easily employed by the typical clinician than other diagnostic options such as tympanometry and acoustic reflectometry. Nevertheless, future studies on the diagnostic assessment of OME should consider cost-effectiveness analysis, which can take into account the variable proficiency of clinicians in performing pneumatic otoscopy as well as the consequences of testing and patient preferences (Sox, Stern, Owen et al., 1989). Cost-effectiveness analysis will enable more informed decisions on the best diagnostic method for OME. The assessment of more complex diagnostic methods such as combination tests or algorithms would also benefit from cost-effectiveness analysis. Such analysis should be undertaken in the future.

References

Abraham SS, Wallace IF, Gravel JS. Early otitis media and phonological development at age 2 years. *Laryngoscope*. 1996;106:727-732.

Alho OP, Koivunen P, Luotonen J, Linder TE, Grenier B. Diagnostic criteria for otitis media in children. *Oto Rhino Laryngologia Nova*. 1998;8:123-128.

Amedee RG. The effects of chronic otitis media with effusion on the measurement of transiently evoked otoacoustic emissions. *Laryngoscope*. 1995;105:589-595.

Aniansson G. Screening diagnosis of secretory otitis media. *Scand Audiol Suppl*. 1986;26:65-69.

Avery CA, Gates GA, Prihoda TJ. Efficacy of acoustic reflectometry in detecting middle ear effusion. *Ann Otol Rhinol Laryngol*. 1986;95:472-476.

Babonis TR, Weir MR, Kelly PC. Impedance tympanometry and acoustic reflectometry at myringotomy [published erratum appears in Pediatrics 1991 Jun;87(6):945]. *Pediatrics*. 1991;87:475-480.

Barnett ED, Klein JO, Hawkins KA, Cabral HJ, Kenna M, Healy G. Comparison of spectral gradient acoustic reflectometry and other diagnostic techniques for detection of middle ear effusion in children with middle ear disease. *Pediatr Infect Dis J*. 1998;17:556-559; discussion 580.

Beery QC, Andrus WS, Bluestone CD, Cantekin EI. Tympanometric pattern classification in relation to middle ear effusions. *Ann Otol Rhinol Laryngol*. 1975;84:56-64.

Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. *Biometrics*. 1999;50:1088-1101.

Ben-David J, Podoshin L, Fradis M. Tympanometry and audiometry in diagnosis of middle-ear effusions. *Ear Nose Throat J*. 1981;60:120-123.

Berman S. Medical management of children with otitis media with effusion. In: Quie PG, ed. *The Report on Pediatric Infectious Diseases*. New York: Churchill Livingstone; 1993:37-38.

Berman S, Roark R, Luckey D. Theoretical cost effectiveness of management options for children with persisting middle ear effusions. *Pediatrics*. 1994;93:353-363.

Birch L, Elbrond O. Prospective epidemiological investigation of secretory otitis media in children attending day-care centers. *ORL J Otorhinolaryngol Relat Spec*. 1984;46:229-234.

Black MM, Gerson LF, Freeland CA, Nair P, Rubin JS, Hutcheson JJ. Language screening for infants prone to otitis media. *J Pediatr Psychol*. 1988;13:423-433.

Black MM, Sonnenschein S. Early exposure to otitis media: a preliminary investigation of behavioral outcome. *J Dev Behav Pediatr*. 1993;14:150-155.

Block SL, Mandel E, McLinn S, et al. Spectral gradient acoustic reflectometry for the detection of middle ear effusion by pediatricians and parents [published erratum appears in Pediatr Infect Dis J 1998 Nov;17(11):1011]. *Pediatr Infect Dis J*. 1998;17:560-564; discussion 580.

Block SL, Pichichero ME, McLinn S, Aronovitz G, Kimball S. Spectral gradient acoustic reflectometry: detection of middle ear effusion in suppurative acute otitis media. *Pediatr Infect Dis J*. 1999;18:741-744.

Bluestone CD, Beery QC, Paradise JL. Audiometry and tympanometry in relation to middle ear effusions in children. *Laryngoscope*. 1973;83:594-604.

Bluestone CD, Cantekin EI. Design factors in the characterization and identification of otitis media and certain related conditions. *Ann Otol Rhinol Laryngol Suppl*. 1979;88:13-28.

Bluestone CD. Definitions, Terminology, and Classification. In: Rosenfeld RM, Bluestone CD, eds. *Evidence-Based Otitis Media*. Saint Louis: B.C. Decker, Inc.; 1999:85-103.

Boswell JB, Nienhuys TG. Reflectometric screening for otitis media: inconsistencies in a sample of Australian aboriginal children. *Int J Pediatr Otorhinolaryngol*. 1993;25:49-60.

Buhrer K, Wall LG, Schuster L. The acoustic reflectometer as a screening device: a comparison. *Ear Hear*. 1985;6:307-314.

Cantekin EI, Beery QC, Bluestone CD. Tympanometric patterns found in middle ear effusions. *Ann Otol Rhinol Laryngol Suppl*. 1977;86:16-20.

Cantekin EI, Bluestone CD, Fria TJ, Stool SE, Beery QC, Sabo DL. Identification of otitis media with effusion in children. *Ann Otol Rhinol Laryngol Suppl*. 1980;89:190-196.

Cantekin EI. Algorithm for diagnosis of otitis media with effusion. *Ann Otol Rhinol Laryngol*. 1983;92:6.

Carlson LH and Stool SE. Diagnosis. In: Rosenfeld RM and Bluestone CD, editors. Evidence-Based Otitis Media. Saint Louis, B.C.: Decker Inc.; 1999. p. 105-116.

Casselbrant ML, Brostoff LM, Cantekin EI, et al. Otitis media with effusion in preschool children. *Laryngoscope*. 1985;95:428-436.

Casselbrant ML, Brostoff LM, Cantekin EI, Ashoff VM, Bluestone CD. Incidence, prevalence, and natural history of otitis media in children in Pittsburgh. *Ann Otol Rhinol Laryngol*. 1990;99:28-29.

Chan LS, Takata GS, Shekelle P, Morton SC, Mason W, Marcy SM. Evidence assessment of management of acute otitis media: II. Research gaps and priorities for future research. *Pediatrics*. (In Press).

Chang SO, Jang YJ, Rhee CK. Effects of middle ear effusion on transient evoked otoacoustic emissions in children. *Auris Nasus Larynx*. 1998;25:243-247.

DerSimonian R, Laird N. Meta-analysis in clinical trial. *Control Clin Trials*. 1986;7:177-188.

Douniadakis DE, Nikolopoulos TP, Tsakanikos MD, Vassiliadis SV, Apostolopoulos NJ. Evaluation of acoustic reflectometry in detecting otitis media in children. *Br J Audiol*. 1993;27:409-414.

Downs MP, Walker DD, Northern JL, Gugenheim S. Identification of Children with Language Delays Due to Recurrent Otitis Media. . *Proceedings of the Fourth International Symposium: Recent Advances in Otitis Media with Effusion*; 1988:382-384.

Duncan RB. Malleus handle probe: a middle ear diagnostic procedure. *Ann Otol Rhinol Laryngol*. 1982;91:281-284.

Egger M, Smith GD, Schneider M, al. e. Bias in meta-analysis detected by a simple, graphical test. *Br Med J*. 1997;315:629-634.

Ernstson S, Sundberg L. Erythromycin in the treatment of otitis media with effusion (OME). *J Laryngol Otol*. 1984;98:767-769.

Feagans L, Sanyal M, Henderson F, Collier A, Appelbaum M. Relationship of middle ear disease in early childhood to later narrative and attention skills. *J Pediatr Psychol*. 1987;12:581-594.

Feagans L, Kipp E, Blood I. The effects of otitis media on the attention skills of day-care-attending toddlers. *Dev Psychol*. 1994;30:701-708.

Feldman HM, Paradise JL, Colborn DK, et al. Word production at age 2 years in relation to otitis media during the first two years of life. . *Proceedings of the Sixth International Symposium: Recent Advances in Otitis Media with Effusion*; 1996:328.

Feldman HM, Dollaghan CA, Campbell TF, et al. Parent-reported language and communication skills at one and two years of age in relation to otitis media in the first two years of life. *Pediatrics*. 1999;104:e52.

Fields MJ, Allison RS, Corwin P, White PS, Doherty J. Microtympanometry, microscopy and tympanometry in evaluating middle ear effusion prior to myringotomy. *N Z Med J*. 1993;106:386-387.

Fiellau-Nikolajsen M. Tympanometry in 3-year-old children. Type of care as an epidemiological factor in secretory otitis media and tubal dysfunction in unselected populations of 3-year-old children. *ORL J Otorhinolaryngol Relat Spec*. 1979;41:193-205.

Fiellau-Nikolajsen M, Lous J. Prospective tympanometry in 3-year-old children. A study of the spontaneous course of tympanometry types in a nonselected population. *Archives of Otolaryngology*. 1979;105:461-466.

Fiellau-Nikolajsen M, Lous J. Tympanometry in three-year-old children. A cohort study on the prognostic value of tympanometry and operative findings in middle ear effusion. *ORL J Otorhinolaryngol Relat Spec*. 1979;41:11-25.

Fiellau-Nikolajsen M. Tympanometry and middle ear effusion: a cohort-study in three-year-old children. *Int J Pediatr Otorhinolaryngol*. 1980;2:39-49.

Fiellau-Nikolajsen M. Tympanometry in three-year-old children. Prevalence and spontaneous course of MEE. *Ann Otol Rhinol Laryngol Suppl*. 1980;89:223-227.

Fiellau-Nikolajsen M. Serial tympanometry and middle ear status in 3-year-old children. *ORL J Otorhinolaryngol Relat Spec*. 1980;42:220-232.

Fiellau-Nikolajsen M, Falbe-Hansen J, Knudstrup P. Tympanometry in three-year-old children. III. Correlation between tympanometry and findings at paracentesis in a prospectively followed population of otherwise healthy children aged 3--4 years. *Scand Audiol*. 1980;9:49-54.

Fiellau-Nikolajsen M. Tympanometry in three-year-old children. The 3-year follow-up of a cohort study. *ORL J Otorhinolaryngol Relat Spec*. 1981;43:89-103.

Fiellau-Nikolajsen M. Epidemiology of secretory otitis media. A descriptive cohort study. *Ann Otol Rhinol Laryngol*. 1983;92:172-177.

Finitzo T, Friel-Patti S, Chinn K, Brown O. Tympanometry and otoscopy prior to myringotomy: issues in diagnosis of otitis media. *Int J Pediatr Otorhinolaryngol*. 1992;24:101-110.

Fischler RS, Todd NW, Feldman CM. Otitis media and language performance in a cohort of Apache Indian children. *Am J Dis Child*. 1985;139:355-360.

Freeark K, Frank SJ, Wagner AE, Lopez M, Olmsted C, Girard R. Otitis media, language development, and parental verbal stimulation. *J Pediatr Psychol*. 1992;17:173-185.

Freyss GE, Narcy PP, Manac'h Y, Toupet MG. Acoustic reflex as a predictor of middle ear effusion. *Ann Otol Rhinol Laryngol Suppl*. 1980;89:196-199.

Fria TJ, Sabo DL. Auditory brainstem responses in children with otitis media with effusion. *Ann Otol Rhinol Laryngol Suppl*. 1980;89:200-206.

Fria TJ, Cantekin EI, Probst G. Validation of an automatic otoadmittance middle ear analyzer. *Ann Otol Rhinol Laryngol*. 1980;89:253-256.

Fried M, Kelly J, Zubick H, Vernick D, Strome M. Acoustic reflectivity for diagnosis of middle ear effusions: Comparison with results of myringotomy. *New Dimensions Otorhinolaryngol- Head Neck Surg*. 1985;2:931-932.

Friel-Patti S, Finitzo-Hieber T, Conti G, Brown KC. Language delay in infants associated with middle ear disease and mild, fluctuating hearing impairment. *Pediatr Infect Dis*. 1982;1:104-109.

Gates GA, Avery CA, Prihoda TJ, Cooper JC, Jr. Effectiveness of adenoidectomy and tympanostomy tubes in the treatment of chronic otitis media with effusion. *N Engl J Med*. 1987;317:1444-1451.

Gates GA. Cost-effectiveness considerations in otitis media treatment. In: Lim D, Bluestone CD, Casselbrant M, Klein J, Ogra P, eds. *Proceedings of the Sixth International Symposium*. Fort Lauderdale, FL: B.C. Decker Inc.; 1996a:1-4.

Gates GA. Cost-effectiveness considerations in otitis media treatment. *Otolaryngol Head Neck Surg*. 1996b;114:525-530.

Gersdorff M, Scholtes JL, Yousif A, Robillard TH. Secretory otitis media, tympanometry and general anesthesia. In: Sade J, ed. *Acute and secretory otitis media*. Amsterdam: Kugler publications; 1986:39-44.

Gravel JS, Wallace IF. Listening and language at 4 years of age: effects of early otitis media. *J Speech Hear Res*. 1992;35:588-595.

Gravel JS, Wallace IF, Ellis MA, Lee WW, Mody M, Ruben RJ. Higher-order auditory abilities at nine years of age and early otitis media. . *Proceedings of the Sixth International Symposium: Recent Advances in Otitis Media with Effusion*; 1996:356-358.

Gravel JS, Wallace IF. Effects of otitis media with effusion on hearing in the first 3 years of life. *J Sp Lang Hear Res*. 2000;43:631-644.

Grimaldi PM. The value of impedance testing in diagnosis of middle ear effusion. *J Laryngol Otol*. 1976;90:141-152.

Haapaniemi JJ. Pure-tone audiometric and impedance measurements in school-aged children in Finland. *Eur Arch Otorhinolaryngol*. 1997;254:269-273.

Hall MJ, Lawrence L. Ambulatory surgery in the Unites States, 1995. Advance data from vital and health statistics; no 296. . Hyattsville (MD): National Center for Health Statistics; 1997.

Harsten G, Nettelbladt U, Schalen L, Kalm O, Prellner K. Language development in children with recurrent acute otitis media during the first three years of life. Follow-up study from birth to seven years of age. *J Laryngol Otol*. 1993;107:407-412.

Haughton PM. Validity of tympanometry for middle ear effusions. *Archives of Otolaryngology*. 1977;103:505-513.

Hedges LV, Olkin I. *Statistical Methods for meta-analysis.* San Diego: Academic Press; 1985.

Holm-Jensen S, Sorensen CH, Tos M. Repetitive tympanometric screenings in 4-year-old children. Seasonal influence on secretory otitis and tubal dysfunction. *ORL J Otorhinolaryngol Relat Spec*. 1981;43:164-174.

Holmquist J, Al Fadala S, Qattan Y. Prevalence of secretory otitis media among school children in Kuwait. *J Laryngol Otol*. 1987;101:116-119.

Hutchings ME, Meyer SE, Moore DR. Binaural masking level differences in infants with and without otitis media with effusion. *Hear Res*. 1992;63:71-78.

Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature: III. How to use an article about diagnostic test: A. Are the results of study valid? *J Am Med Assoc*. 1994;271:389-391.

Johnson LP, Parkin JL, Stevens MH, et al. Action of general anesthesia on middle ear effusions. *Archives of Otolaryngology*. 1980;106:100-102.

Jonathan DA. Sonotubometry: its role in childhood glue ear. *Clinical Otolaryngology and Allied Sciences*. 1989;14:151-154.

Jung TTK. Classification of otitis media and surgical principles. *Otolaryngol Clin North Am.* 1999 Jun;32:369-383.

Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *J Am Med Assoc*. 1999;282:1054-1060.

Kaleida PH, Stool SE, Karabin JC. Assessment of otoscopic accuracy in diagnosis of middle-ear effusion in children ("otoscopic validation"). . *Proceedings of the Fifth International Symposium: Recent Advances in Otitis Media with Effusion*; 1992:41.

Kaplan GJ, Fleshman JK, Bender TR, Baum C, Clark PS. Long-term effects of otitis media: a ten-year cohort study of Alaskan Eskimo children. *Pediatrics*. 1973;52:577-85.

Karma PH, Penttila MA, Sipila MM, Kataja MJ. Otoscopic diagnosis of middle ear effusion in acute and non-acute otitis media. I. The value of different otoscopic findings. *Int J Pediatr Otorhinolaryngol*. 1989;17:37-49.

Kemaloglu YK, Sener T, Beder L, Bayazit Y, Goksu N. Predictive value of acoustic reflectometry (angle and reflectivity) and tympanometry. *Int J Pediatr Otorhinolaryngol*. 1999;48:137-142.

Kennedy TL, Gore LB. Middle ear effusions and the nitrous oxide myth. *Laryngoscope*. 1982;92:169-172.

Klein JO, Chase C, Teele DW, Menyuk P, Rosner BA, Group tGBOMS. Otitis media and the developoment of speech, language and cognititve abilities at seven years of age. In: Lim DJ BC, Klein JO, Nelson JD, ed. *Recent advances in otitis media with effusion*. Philadelphia: Decker; 1988:396-397.

Klein JO, Bluestone CD. Management of otitis media in the era of managed care. *Adv Pediatr Infect Dis*. 1996;12:351-386.

Knishkowy B, Palti H, Adler B, Tepper D. Effect of otitis media on development: a community-based study. *Early Hum Dev*. 1991;26:101-111.

Koivunen P, Alho OP, Uhari M, Niemela M, Luotonen J. Minitympanometry in detecting middle ear fluid. *J Pediatr*. 1997;131:419-422.

Kozak LJ, Hall MJ, Pokras R, Lawrence L. Ambulatory Surgery in the Unites States, 1994. Advance Data from vital and health statistics. . Hyattsville (MD): National Center for Health Statistics; 1997 Mar.

Laird NM, Mosteller F. Some statistical methods for combining experimental results. *Int J Technol Assess Health Care.* 1990;6:5-30.

Lamothe A, Boudreault V, Blanchette M, Tetreault L, Poliquin J. Serous otitis media: a six week prospective study. *J Otolaryngol*. 1981;10:371-379.

Leiberman A, Bartal N. Untreated persistent middle ear effusion. *J Laryngol Otol*. 1986;100:875-878.

Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*. 1999;282:1061-1066.

Lous J, Fiellau-Nikolajsen M. Epidemiology and middle ear effusion and tubal dysfunction. A one-year prospective study comprising monthly tympanometry in 387 non-selected 7-year-old children. *Int J Pediatr Otorhinolaryngol*. 1981;3:303-317.

Lous J, Fiellau-Nikolajsen M. Early otitis media with effusion and reading achievement. *Allgemeinmedizin*. 1988;17:81-84.

Lovette S. Improved diagnosis of middle ear effusions employing the objective otoscope. *Ann Otol Rhinol Laryngol*. 1976;85:229-233.

Luloff AK, Menyk P, Teele DW. Effect of persistent otitis media on the speech sound repertoire of infants. In: Lim DJ BC, Klein JO, Nelson JK, Ogra P, ed. *Recent Advances in Otitis Media: Proceedings of the Fifth International Symposium*: Decker Periodicals; 1993:531-533.

Macknin ML, Skibinski C, Beck G, Hughes G, Kinney S. Acoustic reflectometry detection of middle ear effusion. *Pediatr Infect Dis J*. 1987;6:866-868.

Madell JR. Impact of otitis media on auditory function. In: Rosenfeld RM, Bluestone CD, eds. *Evidence-Based Otitis Media*. Saint Louis: B.C. Decker, Inc.; 1999:337-351.

Mains BT, Toner JG. Pneumatic otoscopy: study of inter-observer variability. *J Laryngol Otol*. 1989;103:1134-1135.

Mandel EM, Rockette HE, Bluestone CD, Paradise JL, Nozza RJ. Myringotomy with and without tympanostomy tubes for chronic otitis media with effusion. *Arch-Otolaryngol-Head-Neck-Surg*. 1989;115:1217-1224.

Mandel EM, Rockette HE, Bluestone CD, Paradise JL, Nozza RJ. Efficacy of myringotomy with and without tympanostomy tubes for chronic otitis media with effusion. *Pediatr Infect Dis J*. 1992;11:270-7.

Marchant CD, McMillan PM, Shurin PA, et al. Objective diagnosis of otitis media in early infancy by tympanometry and ipsilateral acoustic reflex thresholds. *J Pediatr*. 1986;109:590-595.

Marchisio P, Principi N, Passali D, et al. Epidemiology and treatment of otitis media with effusion in children in the first year of primary school. *Acta-Otolaryngol-Stockh*. 1998;118:557-562.

Mathew T, Nordstrom K. On the equivalence of meta-analysis using literature and using individual patient data. *Biometrics*. 1999;55:1221-1223.

Maw AR, Bawden R. Factors affecting resolution of otitis media with effusion in children. *Clin-Otolaryngol*. 1994;19:125-130.

Mills R, Vaughan-Jones R. A prospective study of otitis media with effusion in adults and children. *Clin Otolaryngol Allied Sci*. 1992;17:271-274.

Mitchell DB, Ford GR, Albert D, Waldron J. Acoustic reflectometry as an aid to the diagnosis of glue ear. *Br J Clin Pract*. 1990;44:557-559.

Mody M, Schwartz RG, Gravel JS, Ruben RJ. Speech perception and verbal memory in children with and without histories of otitis media. *J Speech Lang Hear Res*. 1999;42:1069-1079.

Moher D, Pham B, Klassen TP, Schulz KF, Berlin JA, Jadad AR, Liberati A. What contributions do languages other than English make on the results of meta-analyses? *J Clinical Epi. 2000;* 53:964-72.

Moller H, Tos M. Point and period prevalence of otitis media with effusion evaluated by daily tympanometry. *J Laryngol Otol*. 1990;104:937-41.

Nozza RJ, Bluestone CD, Kardatzke D, Bachman R. Towards the validation of aural acoustic immittance measures for diagnosis of middle ear effusion in children. *Ear Hear*. 1992;13:442-453.

Nozza RJ, Bluestone CD, Kardatzke D, Bachman R. Identification of middle ear effusion by aural acoustic admittance and otoscopy. *Ear Hear*. 1994;15:310-323.

Nozza RJ. The Assessment of Hearing and Middle-Ear Function in Children. In: Bluestone CD, Stool SE, and Kenna MA, editors. Pediatric otolaryngology. Philadelphia: W.B. Saunders Company. 1996. p. 165-206.

Nozza RJ, Sabo DL, Mandel EM. A role for otoacoustic emissions in screening for hearing impairment and middle ear disorders in school-age children. *Ear Hear*. 1997;18:227-239.

Orchik DJ, Dunn JW, McNutt L. Tympanometry as a predictor of middle ear effusion. *Arch Otolaryngol*. 1978;104:4-6.

Orchik DJ, Morff R, Dunn JW. Impedance audiometry in serous otitis media. *Arch Otolaryngol*. 1978;104:409-412.

Orchik DJ, Morff R, Dunn JW. Middle ear status at myringotomy and its relationship to middle ear immitance measurements. *Ear Hear*. 1980;1:324-328.

Ovesen T, Paaske PB, Elbrond O. Accuracy of an automatic impedance apparatus in a population with secretory otitis media: principles in the evaluation of tympanometrical findings. *Am J Otolaryngol*. 1993;14:100-104.

Owen MJ, Johnson DL, Swank PR, Baldwin CD, Aker J, Howie VM. Duration of otitis media with effusion from birth to 3 years of age related to language and cognititve scores at 5 years. In: Lim DJ BC, Casselbrant M, Klein JO, Ogra P, ed. *Recent Advances in Otitis Media: Proceedings of the Sixth International Symposium*. Hamilton, Ontario: B.C. Decker Inc; 1996:329-331.

Oyiborhoro JM, Olaniyan SO, Newman CW, Balakrishnan SL. Efficacy of acoustic otoscope in detecting middle ear effusion in children. *Laryngoscope*. 1987;97:495-498.

Palmu A, Puhakka H, Rahko T, Takala AK. Diagnostic value of tympanometry in infants in clinical practice. *Int J Pediatr Otorhinolaryngol*. 1999;49:207-213.

Paradise JL, Smith CG, Bluestone CD. Tympanometric detection of middle ear effusion in infants and young children. *Pediatrics*. 1976;58:198-210.

Paradise JL, Bluestone CD, Rogers KD, 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;263:2066-73.

Paradise JL. Treatment guidelines for otitis media: the need for breadth and flexibility. *Pediatr Infect Dis J*. 1995;14:429-35.

Paradise JL, Smith CG, Sabo DL, et al. Tympanometric detection of middle ear effusion in the first versus the second year of life. . *Proceedings of the Sixth International Symposium: Recent Advances in Otitis Media with Effusion*; 1996:155.

Paradise JL, Rockette HE, Colborn DK, et al. Otitis media in 2253 Pittsburgh-area infants: prevalence and risk factors during the first two years of life [see comments]. *Pediatrics*. 1997;99:318-333.

Paradise JL. Otitis media and child development: should we worry? *Pediatr Infect Dis J*. 1998;17:1076-1083; discussion 1099-1100.

Paradise JL, Dollaghan CA, Campbell TF, et al. Language, speech sound production, and cognition in three-year-old children in relation to otitis media in their first three years of life. *Pediatrics*. 2000;105:1119-1129.

Park IY, Kim HN, Chung MH, Park KH, Yim SB. Diagnostic Reliability of Impedance Audiometry in Case of Otitis Media with Effusion. . *Proceedings of the Fourth International Symposium: Recent Advances in Otitis Media with Effusion*; 1988:52-54.

Patterson M, Paparella MM. Otitis media with effusion and early sequelae. *Otolaryngol Clin North Am*. 1999 Jun;32:391-400.

Paul R, Lynn TF, Lohr-Flanders M. History of middle ear involvement and speech/language development in late talkers. *J Speech Hear Res*. 1993;36:1055-1062.

Pellett FS, Cox LC, MacDonald CB. Use of acoustic reflectometry in the detection of middle ear effusion. *J Am Acad Audiol*. 1997;8:181-187.

Petinou K, Mody M, Schwartz RD, Gravel JS, Ellis MA, Lee WW. Impact of otitis media on early speech development: phonetic inventories. . *Proceedings of the Sixth International Symposium: Recent Advances in Otitis Media with Effusion*; 1996:408-410.

Petinou K, Schwartz RG, Mody M, Gravel JS. The impact of otitis media with effusion on early phonetic inventories: A longitudinal prospective investigation. *Clin Linguistics Phonetics*. 1999;13:351-367.

Portoian-Shuhaiber S, Cullinan TR. Middle ear disease assessed by impedance in primary school children in south London. *Lancet*. 1984;1:1111-1112.

Poulsen G, Tos M. Screening tympanometry in newborn infants and during the first six months of life. *Scand Audiol*. 1978;7:159-166.

Poulsen G, Tos M. Repetitive tympanometric screenings of two-year-old children. *Scand Audiol*. 1980;9:21-28.

Rahko T, Laitila P, Sipila M, Manninen M, Karma P. Hearing and acute otitis media in 13-year-old children. *Acta Otolaryngol*. 1995;115:190-192.

Reed D, Struve S, Maynard JE. Otitis media and hearing deficiency among Eskimo children: a cohort study. *Am J Public Health Nations Health*. 1967;57:1657-1662.

Rees GL, Freeland AP. The effect of anaesthesia on tympanograms of children undergoing grommet insertion [see comments]. *Clin Otolaryngol Allied Sci*. 1992;17:200-202.

Renvall U, Liden G, Jungert S, Nilsson E. Long-term observation of ears with reduced middle ear pressure. *Acta Otolaryngol*. 1978;86:104-9.

Renvall U, Aniansson G, Liden G. Spontaneous improvement in ears with middle ear disease. *Int J Pediatr Otorhinolaryngol*. 1982;4:245-250.

Renvall U, Berg U. Reliability of tympanogram. . *Proceedings of the Sixth International Symposium: Recent Advances in Otitis Media with Effusion*; 1996:143-146.

Reves R, Budgett R, Miller D, Wadsworth J, Haines A. Study of middle ear disease using tympanometry in general practice. *Br Med J Clin Res Ed.* 1985;290:1953-1956.

Roberts JE, Sanyal MA, Burchinal MR, Collier AM, Ramey CT, Henderson FW. Otitis media in early childhood and its relationship to later verbal and academic performance. *Pediatrics*. 1986;78:423-430.

Roberts JE, Burchinal MR, Koch MA, Collier AM, Henderson FW. Otitis Media in Early Childhood and Its Relationship to Later Speech and Language. . *Proceedings of the Fourth International Symposium: Recent Advances in Otitis Media with Effusion*; 1988:385-388.

Roberts JE, Bruchinal MR, Koch MA, Footo MM, Henderson FW. Otitis media in early childhood and its relationship to later phonological development. *J Speech Hear Dis*. 1988;53:416-424.

Roberts JE, Burchinal MR, Collier AM, Ramey CT, Koch MA, Henderson FW. Otitis meida in early childhood and cognititive, academic, and classroom performance of the school-aged child. *Pediatrics*. 1989;83:477-485.

Roberts JE, Burchinal MR, Davis BP, Collier AM, Henderson FW. Otitis media in early childhood and later language. *J Speech Hear Res*. 1991;34:1158-1168.

Roberts DG, Johnson CE, Carlin SA, Turczyk V, Karnuta MA, Yaffee K. Resolution of middle ear effusion in newborns. *Arch Pediatr Adolesc Med*. 1995;149:873-877.

Roberts JE, Burchinal MR, Clarke-Klein SM. Otitis media in early childhood and cognitive, academic, and behavior outcomes at 12 years of age. *J Pediatr Psychol*. 1995;20:645-660.

Roberts JE, Burchinal MR, Medley LP, et al. Otitis media, hearing sensitivity, and maternal responsiveness in relation to language during infancy. *J Pediatr*. 1995;126:481-489.

Roberts JE, Zeisel S, Medley L, et al. Otitis media, hearing sensitivity, and language development to two-year-olds. . *Proceedings of the Sixth International Symposium: Recent Advances in Otitis Media with Effusion*; 1996:325-328.

Roberts JE, Burchinal MR, Zeisel SA, et al. Otitis media, the caregiving environment, and language and cognitive outcomes at 2 years. *Pediatrics*. 1998;102:346-354.

Roberts JE, Burchinal MR, Jackson SC, et al. Otitis media in early childhood in relation to preschool language and school readiness skills among African American children. *Pediatrics.* 2000; 106:1-11.

Roberts JE, Wallace IF. Language and otitis media. In: Roberts JE, Wallace IF, Henderson FW, eds. *Otitis media in young children: medical, developmental and educational considerations*. Baltimore (MD): Brooks; 1997:133-161.

Robinson DO, Allen DV, Root LP. Infant tympanometry: differential results by race. *J Speech Hear Disord*. 1988;53:341-346.

Rosenfeld RM, Mandel EM, Bluestone CD. Systemic steroids for otitis media with effusion in children [see comments]. *Arch Otolaryngol Head Neck Surg*. 1991;117:984-989.

Rosenfeld RM, Post JC. Meta-analysis of antibiotics for the treatment of otitis media with effusion. *Otolaryngol Head Neck Surg*. 1992;106:378-386.

Rosenfeld RM. Comprehensive management of otitis media with effusion. *Otolaryngol Clin North Am*. 1994;27:443-455.

Rosenfeld RM. Natural history of untreated otitis media. In: Rosenfeld RM, Bluestone CD, eds. *Evidence-Based Otitis Media*. Saint Louis: B.C. Decker Inc.; 1999a:157-177.

Rosenfeld RM. What to expect from medical therapy. In: Rosenfeld RM, Bluestone CD, eds. *Evidence-Based Otitis Media*. Saint Louis: B.C. Decker Inc.; 1999b:179-205.

Ruben RJ, Wallace IF, Gravel J. Long-term communication deficiencies in children with otitis media during their first year of life. *Acta Otolaryngol*. 1997;117:206-207.

Ryding M, Konradsson K, Kalm O, Prellner K. Sequelae of recurrent acute otitis media. Ten-year follow-up of a prospectively studied cohort of children. *Acta Paediatr*. 1997;86:1208-1213.

Sacket DL. How to read clinical journals. V: To distinguish useful from useless or even harmful therapy. *Can Med Assoc J*. 1981;124:1156-1162.

Sassen ML, van Aarem A, Grote JJ. Validity of tympanometry in the diagnosis of middle ear effusion. *Clin Otolaryngol Allied Sci*. 1994;19:185-189.

Schappert SM. Office visits for otitis media: United States, 1975-90. Advance data from vital and health statistics. . Hyattsville (MD): National Center for Health Statistics; 1992.

Schappert SM. National ambulatory medical care survey: 1994 summary. Advance data from vital and health statistics. . Hyattsville (MD): National Center for Health Statistics; 1996.

Schappert SM. Ambulatory care visits to physician offices, hospital outpatient departments, and emergency departments: United States, 1995. . Hyattsville (MD): U.S. Department of Health and Human Services; 1997.

Schwartz DM, Schwartz RH. Validity of acoustic reflectometry in detecting middle ear effusion. *Pediatrics*. 1987;79:739-742.

Shaw JO, Stark EW, Gannaway SD. The influence of nitrous oxide anaesthetic on middle-ear fluid. *J Laryngol Otol*. 1978;92:131-135.

Silman S, Silverman CA, Arick DS. Acoustic-immittance screening for detection of middle-ear effusion in children. *J Am Acad Audiol*. 1992;3:262-268.

Silman S, Silverman CA, Arick DS. Pure-tone assessment and screening of children with middle-ear effusion. *J Am Acad Audiol*. 1994;5:173-182.

Silva AB, Hotaling AJ. A protocol for otolaryngology-head and neck resident training in pneumatic otoscopy. *Int J Pediatr Otorhinolaryngol*. 1997;40:125-131.

Silverman CA, Silman S. Acoustic-immittance characteristics of children with middle-ear effusion: longitudinal investigation. *J Am Acad Audiol*. 1995;6:339-345.

Sly RM, Zambie MF, Fernandes DA, Frazer M. Tympanometry in kindergarten children. *Annals of Allergy*. 1980;44:1-7.

Sorri M, Maki-Torkko E, Alho OP. Otitis media and long-term follow-up of hearing. *Acta Otolaryngol Stockh*. 1995;109:811-816.

Sox H, Stern S, Owens D, Abrams HL. Assessment of diagnostic technology in health care: Rationale, methods, problems, and directions. . Washington, D.C.: National Academy Press; 1989.

StataCorp. Stata Statistical Software: Release 6.0. : College Station, TX: Stata Corporation; 1999.

Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet*. 1993;341:418-422.

Stewart LA, Clarke MJ. Practical methodology of meta-analyses (overviews) using updated individual patient data. Cochrane Working Group. *Stat Med*. 1995;14:2057-2079.

Stool SE, Field MJ. The impact of otitis media. *Pediatr Infect Dis J*. 1989;8:S11-S14.

Stool SE, Berg AO, Berman S, al. e. Otitis media with effusion in young children. Clinical Practice Guideline, Number 12. . Rockville (MD): Agency for Health Care Policy and Research, Public Health

Service, U.S. Department of Health and Human Services; 1994a July.

Stool SE, Berg AO, Berman S, al. e. Managing otitis media with effusion in ytoung children. Quick Reference Guide for Clinicians. . Rockville (MD): Agency for Health Care Policy and Research, Public Health Servies, U.S. Department fo Health and Human Services; 1994b July.

Szucs E, Diependaele R, Clement PA. The accuracy of tympanometry assessed by its sensitivity and specificity. *Acta Otorhinolaryngol Belg*. 1995;49:287-292.

Takahashi H, Honjo I, Hasebe S, Sudo M, Tanabe M. The diagnostic and prognostic value of eardrum mobility in otitis media with effusion. *Eur Arch Otorhinolaryngol*. 1999;256:189-191.

Takata GS, Chan LS, Shekelle P, Morton SC, Mason W, Marcy SM. Evidence assessment of management of acute otitis media: I. The role of antibiotics in treatment of uncomplicated acute otitis media. *Pediatrics*. (In Press).

Teele DW, Klein JO, Rosner BA. Epidemiology of otitis media in children. *Ann Otol Rhinol Laryngol Suppl*. 1980;89:5-6.

Teele DW, Klein JO, Rosner BA. Otitis media with effusion during the first three years of life and development of speech and language. *Pediatrics*. 1984;74:282-287.

Teele DW, Klein JO, Chase C, Menyuk P, Rosner BA. Otitis media in infancy and intellectual ability, school achievement, speech, and language at age 7 years. Greater Boston Otitis Media Study Group. *J Infect Dis*. 1990;162:685-694.

Thomsen J, Tos M. Spontaneous improvement of secretory otitis. A long-term study. *Acta Otolaryngol*. 1981;92:493-499.

Tom LW, Tsao F, Marsh RR, Kessler A, Konkle DF. Effect of anesthetic gas on middle ear fluid. *Laryngoscope*. 1994;104:832-836.

Toner JG, Mains B. Pneumatic otoscopy and tympanometry in the detection of middle ear effusion. *Clin Otolaryngol Allied Sci*. 1990;15:121-123.

Tos M, Poulsen G. Tympanometry in 2-year-old children. Seasonal influence on frequency of secretory otitis and tubal function. *ORL J Otorhinolaryngol Relat Spec*. 1979;41:1-10.

Tos M, Poulsen G, Hancke AB. Screening tympanometry during the first year of life. *Acta Otolaryngol*. 1979;88:388-94.

Tos M. Spontaneous improvement of secretory otitis and impedance screening. *Arch Otolaryngol*. 1980;106:345-349.

Tos M, Poulsen G. Screening tympanometry in infants and two-year-old children. *Ann Otol Rhinol Laryngol Suppl*. 1980;89:217-222.

Tos M, Holm-Jensen S, Sorensen CH. Changes in prevalence of secretory otitis from summer to winter in four-year-old children. *Am J Otol*. 1981;2:324-327.

Tos M, Holm-Jensen S, Sorensen CH, Mogensen C. Spontaneous course and frequency of secretory otitis in 4-year-old children. *Arch-Otolaryngol*. 1982;108:4-10.

Tos M. Epidemiology and spontaneous improvement of secretory otitis. *Acta Otorhinolaryngol Belg*. 1983;37:31-43.

Tos M. Epidemiology and natural history of secretory otitis. *Am J Otol*. 1984;5:459-462.

Tos M, Stangerup SE, Andreassen UK, Hvid G, Thomsen J, Holm-Jensen S. Natural History of Secretory Otitis Media. . *Proceedings of the Third International Symposium: Recent Advances in Otitis Media with Effusion*; 1984:36-40.

Tos M, Stangerup SE, Hvid SG, Andreassen UK. Epidemiology and Natural History of Secretory Otitis. . *Proceedings of the Fourth International Symposium: Recent Advances in Otitis Media with Effusion*; 1988:29-34.

Trout KS. How to read clinical journal: IV. To determine etiology or causation. *Can Med Assoc J*. 1981;124:985-990.

Tsushima T, Menyuk P, Teele DW. Effects of otitis media on syllable realization in the speech of children. In: Lim DJ BC, Klein JO, Nelson JD, Ogra P, ed. *Recent Advances in Otitis Media: Proceedings of the Fifth International Symposium*: Decker Periodicals; 1993:573-575.

Tugwell P. How to read clinical journals: III. To learn the clinical course and prognosis of disease. *Can Med Assoc J*. 1981;124:869-872.

U.S. BOTC. Statistical abstract of the United States, 1998. . Springfield (VA): National Technical Information Services; 1996.

U.S. BOTC. Statistical abstract of the United States, 1998. . Springfield (VA): National Technical Information Services; 1998.

van Balen FA, de Melker RA. Validation of a portable tympanometer for use in primary care. *Int J Pediatr Otorhinolaryngol*. 1994;29:219-225.

van Balen FA, de Melker RA, Touw-Otten FW. Double-blind randomised trial of co-amoxiclav versus placebo for persistent otitis media with effusion in general practice [see comments]. *Lancet*. 1996;348:713-716.

Vaughan-Jones R, Mills RP. The Welch Allyn Audioscope and Microtymp: their accuracy and that of pneumatic otoscopy, tympanometry and pure tone audiometry as predictors of otitis media with effusion. *J Laryngol Otol*. 1992;106:600-602.

Vernon-Feagans L, Manlove EE, Volling BL. Otitis media and the social behavior of day-care-attending children. *Child Dev*. 1996;67:1528-1539.

Vernon-Feagans L, Emanuel D, Blood I. Otitis media and day care: factors hat affect children's language. . *Proceedings of the Sixth International Symposium: Recent Advances in Otitis Media with Effusion*; 1996:331-334.

Vernon-Feagans L. Impact of Otitis Media on Speech, Language, Cognition, and Behavior. In: Rosenfeld RM, Bluestone CD, eds. *Evidence-Based Otitis Media*. Saint Louis: B.C. Decker Inc.; 1999:353-373.

Wallace IF, Gravel JS, McCarton CM, Ruben RJ. Otitis media and language development at 1 year of age. *J Speech Hear Disord*. 1988;53:245-251.

Wallace IF, Gravel JS, McCarton CM, Stapells DR, Bernstein RS, Ruben RJ. Otitis Media and Its Effect on Language in Infancy. . *Proceedings of the Fourth International Symposium: Recent Advances in Otitis Media with Effusion*; 1988:388-391.

Wallace IF, Gravel JS, McCarton CM, Stapells DR, Bernstein RS, Ruben RJ. Otitis media, auditory sensitivity, and language outcomes at one year. *Laryngoscope*. 1988;98:64-70.

Wallace IF, Gravel JS, Ganon EC, Ruben RJ. Two-year language outcomes as a function of otitis media and parental linguistic styles. . *Proceedings of the Fifth International Symposium: Recent Advances in Otitis Media with Effusion*; 1992:527-530.

Wallace IF, Gravel JS, Schwartz RG, Ruben RJ. Otitis media, communication style of primary caregivers, and language skills of 2 year olds: a preliminary report. *J Dev Behav Pediatr*. 1996;17:27-35.

Watters GW, Jones JE, Freeland AP. The predictive value of tympanometry in the diagnosis of middle ear effusion. *Clin Otolaryngol Allied Sci*. 1997;22:343-345.

Webster A, Bamford JM, Thyer NJ, Ayles R. The psychological, educational and auditory sequelae of early, persistent secretory otitis media. *J Child Psychol Psychiatry*. 1989;30:529-546.

Wendler-Shaw PD, Menyuk P, Teele DW. Effects of otitis media in the first year of life on language production in the second year of life. In: Lim DJ BC, Klein JO, Nelson JK, Ogra P, ed. *Recent Advances in Otitis Media: Proceedings of the Fifith International Syposium*. Hamilton, Ontario: Decker Periodicals; 1993:534-536.

Williams RG, Haughton PM. Tympanometric diagnosis of middle ear effusions. *J Laryngol Otol*. 1977;91:959-962.

Williams RL, Chalmers TC, Stange KC, Chalmers FT, Bowlin SJ. Use of antibiotics in preventing recurrent acute otitis media and in treating otitis media with effusion. A meta-analytic attempt to resolve the brouhaha [published erratum appears in JAMA 1994 Feb 9;271(6):430] [see comments]. *JAMA*. 1993;270:1344-1351.

Williamson IG, Dunleavey J, Bain J, Robinson D. The natural history of otitis media with effusion--a three-year study of the incidence and prevalence of abnormal tympanograms in four South West Hampshire infant and first schools. *J Laryngol Otol*. 1994;108:930-934.

Wilmot JF, Cable HR. Persistent effusion following acute otitis media: tympanometry and pneumatic otoscopy in diagnosis. *J R Coll Genl Practitioners*. 1988;38:149-152.

Woodwell DA, Schappert SM. National ambulatory medical care survey: 1993 summary. Advance data from vital and health statistics. . Hyattsville (MD): National Center for Health Statistics; 1995.

Woodwell DA. National ambulatory medical care survey: 1995 summary. Advance data from vital and health statistics. . Hyattsville (MD): National Center for Health Statistics; 1997a.

Woodwell DA. National ambulatory medical care survey. Advance data from vital and health statistics. . Hyattsville (MD): National Center for Health Statistics; 1997b.

Wright P, Thompson J, McConnell KB, Sitton AB, Bess FH. Audiologic and Speech Evaluation of a Prospectively Followed Cohort of Normal Children: The Impact of Otitis Media. . *Proceedings of the Third International Symposium: Recent Advances in Otitis Media with Effusion*; 1984:338-341.

Wright PF, Sell SH, McConnell KB, et al. Impact of recurrent otitis media on middle ear function, hearing, and language. *J Pediatr*. 1988;113:581-587.

Zargi M, Boltezar IH. Effects of recurrent otitis media in infancy on auditory perception and speech. *Am J Otolaryngol*. 1992;13:366-372.

Zeisel SA, Roberts JE, Gunn EB, et al. Prospective surveillance for otitis media with effusion among black infants in group child care. *J Pediatr*. 1995;127:875-880.

Zielhuis GA, Rach GH, van den Broek P. The natural course of otitis media with effusion in preschool children. *Eur Arch Otorhinolaryngol*. 1990;247:215-221.

Evidence Table 1. Natural History of Otitis Media with Effusion

[The six components of study quality are: a prospective cohort study; outcome clearly defined; time point at which outcome measured clearly defined; subjects followed without any intervention; blinded assessment of outcome; and point and variability estimates provided for main outcome measures. 1 indicates presence and 0 indicates absence. OME resolution rates, not otherwise specified, represent the proportion without OME at follow-up compared to those with OME at baseline.]

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
862 Birch and Elbrønd 1984	Study Type: prospective single cohort Study Quality Score (0–6): 3(111000) OME definition: type B tympanogram signified middle-ear effusion otherwise not defined Group: Children attending a day-care center in a municipality under investigation for incidence of OME N=373 subjects	Time: Jan.–Apr./1982 Place: 10 day-care centers in a municipality of about 20,000 inhabitants; Dept. of Oto-Rhino-Laryngology, University of Aarhus, Denmark Inclusion: • attending day-care centers in investigated municipality Exclusion: • tympanometry not practicable (n=15 children) • parental consent not given (n=5 children) • ill or absent at all examinations (n=3 children)	Type of OME: Newly diagnosed OME of unknown duration; OME persisting for weeks or months Age: 0.75–7 years, mean 4.5 years Examiner(s): unknown OME diagnostic method: a B tympanogram defined as “compliance below 0.25 ml” signified middle-ear effusion Interventions: The authors did not mention control for interventions such as antibiotics during the course of the study, though the number receiving surgical procedures was noted. Interval of screening: every 2 weeks for twelve weeks	Data not abstractable. The authors present an equation to estimate OME resolution, most likely cumulative rates, reportedly based on the study data. $Y=Y_0e^{-kt}$ Y= number of ears with OME Y_0=number of ears with OME a time 0 t=time in weeks k=0.30 Derived OME resolution rates by ear Week — % remain — % resolved 0 — 100 — 0 1 — 74 — 26 2 — 55 — 45 3 — 41 — 59 4 — 30 — 70 5 — 22 — 78 6 — 17 — 83 7 — 12 — 88 8 — 9 — 91 9 — 7 — 93 10 — 5 — 95 11 — 4 — 96 12 — 3 — 97 13 — 2 — 98 14 — 1 — 99 15 — 1 — 99 16 — 1 — 99 17 — 1 — 99 18 — 0 — 100

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
1000 Casselbrant, Brostoff, Cantekin, Doyle, Bluestone, and Fria 1985 [This cohort is also in article 2929.]	Study Type: prospective comparative cohort Study Quality Score (0–6): 3(111000) OME definition: inflammation of the middle ear accompanied by collection of liquid in the middle ear cleft without the signs and symptoms of acute infection such as otalgia, fever, and a red or white tympanic membrane that is full or bulging Group: Children, 2–5 years of age, attending a nursery school and examined from Sep 1981 through Aug 1983 N=103 [103 children were entered into the study, 66 in the first year and 37 in the second year. 37 of the 66 children entered in the first year were also observed during the second year. Therefore in data analysis, the 103 children were treated as 140 'individuals' to account for the 37 children who were observed for both years.]	Time: ~9/1981–8/1983 Place: A nursery school in Verona, a suburb of Pittsburgh Inclusion: • 2–5 years old • nursery school in Verona, a suburb of Pittsburgh Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: 2 year old 23, 3 year old 25, 4 year old 38, 5 year old 17 Gender: male 81, female 59 (37 children in years 1 and 2 were counted twice by the authors) Race/ethnicity: black 4, oriental 1, white 98 Examiner(s): otolaryngologist, nurse trained in tympanometry, and an audiologist; also, validated otoscopists OME diagnostic method: an algorithm combining pneumatic otoscopy, tympanometry, and acoustic reflex measurements Interventions: The authors informed parents if OME was present 3 consecutive months. 11 children received tympanostomy tubes, and 26 children were treated for 32 episodes of acute otitis media of which 13 were associated with a new episode of OME. Interval of screening: monthly in school and every 4 months at the Children's Hospital of Pittsburgh for 2 years after the initial evaluation	Cumulative OME resolution by **episode** (Figure 3) Resolution Time # resolved/# at risk (%) <1 month 92/137 (67.2%) <2 months 109/137 (79.6%) <3 months 130/137 (94.9%) <4 months 134/137 (97.8%) <5 months 136/137 (99.3%) <6 months 137/137 (100%) [This table only includes episodes of known duration without missing exam either prior to onset, during the episode, at resolution, or tympanostomy tubes inserted.]

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings[a]
2929 Casselbrant, Brostoff, Cantekin, Ashoff, and Bluestone 1990 [The 2–5-year-old cohort is also in article 1000.]	Study Type: prospective comparative cohort Study Quality Score (0–6): 3(111000) OME definition: not defined[a] Group 1: Children, 2–5 years of age, attending a nursery school and examined from Sep 1981 through Aug 1983[a] Group 2: School children, 5–12 years of age, attending an elementary school and examined from Sep 1984–May 1985 N=214 N1=103[a] N2=111	Time: Group1~9/1981–8/1983[a]; Group2~9/1984–5/1985 Place: A nursery school in Verona, a suburb of Pittsburgh, [a] and Falk Elementary School, Pittsburgh, PA Inclusion: • Group 1: 2–5 years old and in Verona nursery school • Group 2: 5–12 years old and in Falk Elementary School Exclusion: None[a]	Type of OME: newly diagnosed OME of unknown duration[a] Age: Group1~2–5 years old[a] Group2~5–12 years old Examiner(s): otolaryngologist and nurse practitioner validated otoscopists, nurse trained in tympanometry, and an audiologist[a] OME diagnostic method: an algorithm combining pneumatic otoscopy, tympanometry, and acoustic reflex measurements[a] Interventions: Intervention was advised if OME persisted after 3 months. Up to three months, the authors report no intervention though how control was monitored is not mentioned. [a] Interval of screening: monthly after the initial evaluation[a]	Cumulative OME resolution by **episode** for both groups combined[b] (page 29) Time at resolution % resolved <1 month ≈ 60% <2 months ≈ 80% [b]Actual counts are not given, just approximate percentage of OME resolution at 1 and 2 months.

[a] See 1000 Casselbrant, Brostoff, Cantekin, Doyle, Bluestone, and Fria (1985).

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
1202 Ernstson and Sundberg 1984	Study Type: prospective single cohort embedded in a controlled trial (not randomized) Study Quality Score (0–6): 2(100100) OME definition: "nonpurulent effusion behind an intact tympanic membrane at otomicroscopy…supported by a type "B" tympanogram" Group 1: A~Untreated children with OME at least 3 months scheduled for myringotomy Group 2: B~Treated children with OME at least 3 months scheduled for myringotomy (erythromycin orally 20–30 mg/kg body weight twice a day 10 days preceding surgery) N=98 subjects N1=72 N2=26	Time: not specified Place: The ENT Department, Central Hospital, Karlskrona, Sweden Inclusion: • OME on one or both sides for at least 3 months Exclusion: • cleft palate; immotile cilia syndrome • received antibiotics during the period of observation	Type of OME: OME persisting for weeks or months, i.e. at least 3 months, laterality not specified Age: 1–11 years Laterality: both unilateral and bilateral, proportions not specified Examiner(s): unknown OME diagnostic method: otomicroscopy tympanometry Interventions: The authors did not mention control for interventions such as antibiotics during the pre-operative period prior to the 10 days before myringotomy. Interval of screening: 2–5 weeks, median 3 weeks, between decision to operate and the operation itself.	OME resolution in Group 1 (A) over 2–5 weeks by **child** (page 768) Resolution Time # resolved/# at risk (%) <5 weeks 11/72 (15.3%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 1237, 1242, and 1245 for other findings]
1235 Fiellau-Nikolajsen 1979 [1976 cohort is also in articles 1237, 1242, and 3051. 1978 cohort is also in article 1245.]	Study Type: two cross-sectional studies Study Quality Score (0–6): [a] OME definition: not defined [a] Group 1: Children aged 36–48 months in January, 1976, living in the municipality of Hjoerring, Denmark Group 2: Children aged 36–48 months in August, 1978, living in the municipality of Hjoerring, Denmark N1=503 children, 1001 ears, tested out of 523 [a] N2=435 children, 867 ears, tested out of 465 [a]	Time: Not specified[a] Place: ENT Clinic, Hjoerring Hospital, Denmark[a] Inclusion: [a] • Age: 3 years • residing in Hjoerring, Denmark • had tympanometric screening and rescreening Exclusion: None[a]	Type of OME: newly diagnosed OME of unknown duration[a] Age: 36–48 months old[a] Gender: Group 1 263 males, 260 females; Group 2 234 males, 229 females[a] Daycare: Group 1 18.5% in public day-care center, 30.0% in private home day-care, 53.5% in own home; Group 2 31.1% in public day-care center, 23.8% in private home day-care, 46.2% in own home[a] Examiner(s): unknown[a] OME diagnostic method: tympanometry • A: > –99 mmH_20 • B: ≤ 0.1 relative gradient • C1: –100 to –199 mmH_2O • C2: <–200 mmH_2O Interventions: The authors did not mention control of interventions. [a] Interval of screening: 1, 3 and 6 months after initial diagnosis. Other articles on the 1976 cohort note that once type A tympanogram was noted; follow-up was discontinued.	Cumulative OME resolution by **ear** for Group 2 1978 cohort with denominator from Aug. 1978 (Tables V and VII) **<1 month resolution–all subjects** Type # resolved/# at risk (%) B to A 7/64 (10.9%) B to A/C1 14/64 (21.9%) B to A/C1/C2 22/64 (34.4%) B/C2 to A 23/107 (21.5%) B/C2 to A/C1 36/107 (33.6%) B/C2/C1 to A 76/203 (37.4%) **<1 month resolution–subjects in Daycare** Type # resolved/# at risk (%) B to A 3/35 (8.6%) B to A/C1 6/35 (17.1%) B to A/C1/C2 12/35 (34.3%) B/C2 to A 11/58 (19.0%) B/C2 to A/C1 16/58 (27.6%) B/C2/C1 to A 28/102 (27.5%) **<1 month resolution–subjects at home** Type # resolved/# at risk (%) B to A 4/29 (13.8%) B to A/C1 8/29 (27.6%) B to A/C1/C2 10/29 (34.5%) B/C2 to A 12/49 (24.5%) B/C2 to A/C1 20/49 (40.8%) B/C2/C1 to A 48/101 (47.5%) *Continued on next page*

[a] See 1237 Fiellau-Nikolajsen and Lous (1979) and 1245 Fiellau-Nikolajsen (1983).

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 1237, 1242, and 1245 for other findings]
Continued 1235 Fiellau-Nikolajsen 1979 [1976 cohort is also in articles 1237, 1242, and 3051. 1978 cohort is also in article 1245.]				***Continued*** **<3 month resolution–all subjects** Type # resolved/# at risk (%) B to A 16/62 (25.8%) B to A/C1 25/62 (40.3%) B to A/C1/C2 35/62 (56.5%) B/C2 to A 31/104 (29.8%) B/C2 to A/C1 47/104 (45.2%) B/C2/C1 to A 83/200 (41.5%) **<3 month resolution–subjects in Daycare** Type # resolved/# at risk (%) B to A 4/34 (11.8%) B to A/C1 10/34 (29.4%) B to A/C1/C2 15/34 (44.1%) B/C2 to A 9/56 (16.1%) B/C2 to A/C1 20/56 (35.7%) B/C2/C1 to A 27/100 (27.0%) **<3 month resolution–subjects at home** Type # resolved/# at risk (%) B to A 12/28 (42.9%) B to A/C1 15/28 (53.6%) B to A/C1/C2 20/28 (71.4%) B/C2 to A 22/48 (45.8%) B/C2 to A/C1 27/48 (56.3%) B/C2/C1 to A 56/100 (56.0%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 1242 for longer term findings]
1237 Fiellau-Nikolajsen and Lous 1979 [This cohort is also in article 1235 1242, and 3051.]	Study Type: prospective single cohort Study Quality Score (0–6): 4(111001) OME definition: none given but tympanogram B and C2 were regarded as abnormal Group: Children 3 years of age, born in 1972, and living in the municipality of Hjoerring, Denmark N=523 children fulfilled criteria; 504 (1005 ears) tested. N for type B at cohort inception=98 ears, 56 males and 14 females. N for type B/C2 at cohort inception=220 ears, 112 males and 108 females. N for type B/C2/C1 at cohort inception=372 ears, 186 males and 186 females.	Time: Jan–July, 1976 Place: Municipality of Hjoerring, Denmark (provincial town with about 20,000 inhabitants) Inclusion: • 3 years old born in 1972 in the municipality of Hjoerring, Denmark Exclusion: during follow up: Moved out of area; airtight fitting of probe in auditory canal proved impossible; fistula developed in the pars flaccida	Type of OME: newly diagnosed OME of unknown duration Age: 3 years old (born in 1972) Examiner(s): unknown OME diagnostic method: tympanometry • A: > –100 mm$H_2$0 and gradient > 0.1 • B: 200 to –400 mmH_2O or indeterminable and ≤ 0.1 gradient • C1: –100 to –199 mmH_2O and >0.1 gradient • C2: –200 to –400 mmH_2O and >0.1 gradient Interventions: The authors state that “surgical procedures were avoided as much as possible.” Six children had adenoidectomy. Interval of screening: 1, 3, and 6 months after the initial diagnosis. Once a child had a type A tympanogram follow-up was discontinued.	Cumulative OME resolution by **ear** (Figure 7) **<1 month resolution–all subjects (Fig 7)** Time # resolved/# at risk (%) B to A 14/94 (14.9%) B to A/C1 20/94 (21.3%) B to A/C1/C2 38/94 (40.4%) B/C2 to A 54/212 (25.5%) B/C2 to A/C1 75/212 (35.4%) B/C2/C1 to A 83/266 (31.2%) **<1 month resolution–males** Time # resolved/# at risk (%) B to A 0/52 (0.0%) B to A/C1 3/52 (21.3%) B to A/C1/C2 17/52 (32.7%) B/C2 to A 18/107 (16.8%) B/C2 to A/C1 30/107 (28.0%) B/C2/C1 to A 26/127 (20.5%) **<1 month resolution–females** Time # resolved/# at risk (%) B to A 14/42 (33.3%) B to A/C1 17/42 (40.5%) B to A/C1/C2 21/42 (50.0%) B/C2 to A 36/105 (34.3%) B/C2 to A/C1 45/105 (42.9%) B/C2/C1 to A 57/139 (41.0%) **<3 month resolution–all subjects (Tab 2)** Time # resolved/# at risk (%) B to A 22/91 (24.2%) B to A/C1 30/91 (33.0%) B/C2 to A 85/204 (41.8%) B/C2 to A/C1 106/204 (51.5%) *Continued on next page*

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 1242 for longer term findings]
Continued 1237 Fiellau-Nikolajsen and Lous 1979 [This cohort is also in article 1235 1242, and 3051.]				*Continued* **<3 month resolution–male subjects** Time # resolved/# at risk (%) B to A 5/49 (10.1%) B to A/C1 9/49 (18.4%) B/C2 to A 33/103 (32.0%) B/C2 to A/C1 46/103 (44.7%) **<3 month resolution–female subjects** Time # resolved/# at risk (%) B to A 17/42 (40.5%) B to A/C1 21/42 (50.0%) B/C2 to A 52/101 (51.5%) B/C2 to A/C 60/101 (59.4%) **<6 month resolution–all subjects** Time # resolved/# at risk (%) B to A 32/91 (35.2%) B to A/C1 41/91 (45.1%) B/C2 to A 120/204 (58.8%) B/C2 to A/C1 136/204 (66.7%) **<6 month resolution–male subjects** Time # resolved/# at risk (%) B to A 12/49 (24.5%) B to A/C1 16/49 (32.7%) B/C2 to A 54/103 (52.4%) B/C2 to A/C1 63/103 (61.1%) **<6 month resolution–female subjects** Time # resolved/# at risk (%) B to A 20/42 (47.6%) B to A/C1 25/42 (59.5%) B/C2 to A 66/101 (65.3%) B/C2 to A/C1 73/101 (72.3%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 1237 for earlier findings]
1242 Fiellau-Nikolajsen 1981 [This cohort is also in article 1237, 1235 and 3051.]	Study Type: prospective single cohort within two cross-sectional studies[a] Study Quality Score (0–6): 4(111001) [a] OME definition: not defined, but tympanogram B and C2 were regarded as abnormal[a] Group 1: Children born in 1972 and domiciled in the city of Hjoerring, Denmark in Jan 1976 Group 2: Children born in 1972 and domiciled in the city of Hjoerring, Denmark in Jan 1979 N1=523; 503 tested (in article 1237 n=504 were reported tested out of 523)[a] N2=505; 498 tested; 352 screened in 1976)[a] N for type B at cohort inception=98 ears. N for type B/C2 at cohort inception=220. N for type B/C1/C2 at cohort inception=372.	Time: 3-year followup (1/1976–1/1979) [a] Place: Testing done in classrooms; the municipality of Hjoerring, Denmark[a] Inclusion: [a] • Age: 3–6 years (cohort born in 1972~screened at ages 3 & 6) • living in Hjoerring, Denmark Exclusion: [a] • showed A tympanograms in both ears at a pminimum of one test	Type of OME: newly diagnosed OME of unknown duration[a] Age: 3–6 years (cohort born in 1972~screened at ages 3 & 6) [a] Examiner(s): unknown[a] OME diagnostic method: tympanometry • A: 200 to –99 mmH_20 and relative gradient > 0.1 • B: 200 to –400 mmH_2O or indeterminable and ≤ 0.1 relative gradient • C1: –100 to –199 mmH_2O and >0.1 relative gradient • C2: –200 to –400 mmH_2O and >0.1 relative gradient Interventions: The authors did not mention control of medications. Some received surgical treatment during the follow-up period; 46 paracentesis and adenoidectomy and later 14 of those myringostomy with insertion of grommets; 10 paracentesis and adenoidectomy and later 4 of those myringostomy[a] Interval of screening: 1, 3, and 6 months after the initial diagnosis at 3 years of age as reported in article 1237. This articles describes follow-up at 6 years of age. Once a child had a type A tympanogram follow-up was discontinued. [a]	Cumulative OME resolution by **ear** with denominators from initial Jan. 1976 exam in Figure 1 article 1237 (Table IIa) **<3 years resolution** Type # resolved/# at risk (%) B to A 33/65 (51%) B to A/C1 43/65 (66%) B to A/C1/C2 50/65 (77%) B/C2 to A 85/153 (56%) B/C2 to A/C1 113/153 (74%) B/C1/C2 to A 142/159 (89%) [The authors of article 1242 did not report on the middle-ear status during the three-year interval between the two screenings but comment that there must have been "periodical spontaneous improvements…" (page 98)] With respect to subgroups, the authors of article 1242 mention that "When tested at 6 years,…the named subgroups (gender, public institutions versus home, and urban versus rural) showed no difference, either with respect to impedance test or tone audiometry, and the prognostic role of the tympanogram type at 3 years was the same in all subgroups." Data are not provided on this question. (page 100)

[a] See 1237 Fiellau-Nikolajsen and Lous (1979).

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 1235 for further findings]
1245 Fiellau-Nikolajsen 1983 [This cohort is also in article 1235.]	Study Type: prospective cohort study Study Quality Score (0–6): 3(111000) OME definition: middle-ear effusion with intact tympanic membrane and absence of symptoms of acute inflammation Group: Children screened (underwent audiometry testing and pneumatic otoscopy four times during study period) N=404 subjects screened; N had OME at cohort inception=82.	Time: 1978–1979 (four screens: Aug, Sep & Nov, 1978, Feb 1979) Place: Subject identified in municipality of Hjoerring (medium-size urban town with a population of approximately 20,000); further testing and surgeries performed at ENT Dept., Hjoerring Hospital, Denmark Inclusion: • Age: 3 years (36–48 months) • living in municipality of Hjoerring, Denmark Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: 3 years (36–48 months) Examiner(s): unknown OME diagnostic method: tympanometry • middle-ear effusion: flat curve, or pressure ≤100 mmH_2O with absent middle ear reflexes • no middle-ear effusion Interventions: The authors did not mention control of medications. None reportedly had adenoidectomy, myringotomy, tubes, or paracentesis during the study period. Interval of screening: 1, 3, and 6 months after initial diagnosis	Cumulative OME resolution by **child** (page 173) Resolution Time # resolved/# at risk (%) <1 month 28/78 (35.9%) <3 months 46/78 (59.0%) <6 months 53/78 (68.0%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 543 for further findings]
1486 Holm-Jensen, Sørenson, and Tos 1981 [This cohort is also in articles 2636, 543, 2639, 2642, 4834, and 4835.]	Study Type: prospective cohort study Study Quality Score (0–6): 3(111000) OME definition: not defined though equated to type B tympanogram Group: Healthy children, 4 years old and born on the first-tenth day of each month in 1975, who resided in two Copenhagen counties N=335 children at the first follow-up exam and 333 children at the second follow-up exam N for type B at cohort inception=102 ears (101 in article 2631) N for type B/C2 at cohort inception=318 ears N for type B/C2/C1 at cohort inception=477 ears	Time: February, May, and August, 1979 Place: two counties in Copenhagen, Denmark Inclusion: • born on the 1st–10th of every month in 1975 in two Copenhagen counties • healthy Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: 4 years Examiner(s): unknown OME diagnostic method: tympanometry • A: 0 to –99 mmH_2O • B: flat curve • C1: –100 to –199 mmH_2O • C2: –200 to –350 mmH_2O Interventions: The authors did not mention control of interventions. Interval of screening: 3 and 6 months after initial diagnosis	OME resolution by **ear** the initial Feb. 1979 denominators (Tables I, III, and V) **<3 months resolution** Type # resolved/# at risk (%) B to A 3/92 (3.3%) B to A/C1 14/92 (15.2%) B to A/C1/C2 54/92 (58.7%) B/C2 to A 53/285 (18.6%) B/C2 to A/C1 113/285 (39.6%) B/C2/C1 to A 113/424 (26.7%) **<6 months resolution** Type # resolved/# at risk (%) B to A 14/93 (15.1%) B to A/C1 25/93 (26.9%) B to A/C1/C2 54/93 (58.1%) B/C2 to A 59/282 (20.9%) B/C2 to A/C1 131/282 (46.5%) B/C2/C1 to A 106/418 (25.4%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
1494 Holmquist, Fadala, and Qattan 1987	Study Type: prospective cohort study Study Quality Score (0–6): 3(111000) OME definition: middle-ear effusion behind an intact tympanic membrane Group: Children in grade two who were examined at baseline and three months later by tympanometry and otoscopy N=817 children, 1,634 ears; 788 examined three months later)	Time: 2/1983–4/1983 Place: Eight primary schools distributed in the four educational areas of Kuwait with one male and one female school selected randomly from each area Inclusion: • primary schools Exclusion: • incomplete data • refused to be submitted to examination • occluding ear canal wax • tympanic membrane perforations	Type of OME: newly diagnosed OME of unknown duration Age: 7–9.5 years Examiner(s): the authors whose roles are not specifically identified, one is an MD/PhD, one a surgeon, and one an MS OME diagnostic method: tympanometry • A: +50 to –99 mmH_2O • B: flat curve • C: +100 to –300 mmH_2O Interventions: The authors did not mention control of interventions. Interval of screening: follow-up 3 months after initial diagnosis	OME resolution by **ear** (page 117) **<3 month resolution** Type # resolved/# at risk (%) B/C to A 251/512 (49.0%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
1714 Lamothe, Boudreault, Blanchette, Tetreault, and Poliquin 1981	Study Type: prospective single cohort Study Quality Score (0–6): 4(111100) OME definition: not defined Group: first-grade children N=958 children of which 68 had OME by otoscopy and pneumatic otoscopy after six week follow-up. N of OME at inception: total=68 target ears, male=29, female=39, left ear=38, and right ear=30.	Time: 10/9/1979–12/15/1979 Place: elementary schools in the Sherbrooke metropolitan area, Canada Inclusion: • First graders attending elementary schools in the Sherbrooke metropolitan area, Canada Target ear at first visit was: A) diagnosed with OME if unilateral, B) most affected if bilateral, C) chosen at random if bilateral and evolution was same in both ears Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: First graders Gender: 29 boys, 39 girls Examiner(s): otolaryngologist performed otoscopy and pneumatic otoscopy; audiologist performed tympanogram and audiogram OME diagnostic method: otoscopy and pneumatic otoscopy establishing • abnormal: air-fluid level, bubbles or complete effusion in the middle ear and hypo- or immobility of the eardrum plus two of the three following signs, eardrum retraction, opaque and thickened eardrum, or slight hyperemia tympanometry • abnormal: <= 200 mmH20 or amplitude <= 3 audiogram • abnormal: threshold >= 25 dB Interventions: The authors specifically state the patients were left untreated. Interval of screening: 3 and 6 weeks after initial diagnosis	OME resolution rate by target **ear** (Figure 2) **<3 week resolution** Subjects # resolved/# at risk (%) Total 24/64 (37.5%) Male 7/27 (25.9%) Female 17/37 (45.9%) Left ear 7/35 (20.0%) Right ear 17/29 (58.6%) **<6 week resolution** Subjects # resolved/# at risk (%) Total 25/53 (47.2%) Male 6/23 (26.1%) Female 20/33 (60.6%) Left ear 10/29 (34.5%) Right ear 16/27 (59.3%) Cumulative OME resolution rate by target **ear** (Figure 2) **<3 week resolution** (Total same as above) **<6 week resolution** Subjects # resolved/# at risk (%) Total 38/53 (71.7%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
1735 Leiberman and Bartal 1986	Study Type: retrospective-prospective single cohort Study Quality Score (0–6): 2 (110000) OME definition: not defined Group: Children with persistent middle ear effusion who had follow-up after 2-1/2 year delay in insertion of ventilating tubes N=79 children, 158 ears	Time: 2 1/2 years (actual dates not specified) Place: The Soroka Medical Center in Beer-Sheva, Israel Inclusion: • children with persistent MEE and scheduled for myringotomy and insertion of ventilating tubes which was delayed 2 1/2 years Exclusion: • otomicroscopic findings of fluid and atelectasis, attic retractions, cholesteatoma, perforations, or tympanosclerosis	Type of OME: OME persisting for weeks or months, i.e. 3 months or greater, laterality not specified Age: 2–12 years, mean 6 years Gender: 1:1 (male: female ratio) Examiner(s): unknown who assessed for OME; certified audiologist performed audiometry OME diagnostic method: not stated Interventions: The authors did not mention control for interventions such as antibiotics during the course of the 2-1/2 year delay in surgical intervention. Interval of screening: follow-up greater than 2.5 years after decision to operate for persistent middle-ear effusion of greater-than 3 months	OME resolution by **ear**-(page 876) Resolution Time # resolved/# at risk (%) 2.5 years 49/158 (31.0%) [Authors noted that the middle-ear effusion noted could be either persistent or recurrent. Follow-up for effusion after such a prolonged period may have little meaning.]

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
1777 Lous and Fiellau-Nikolajsen 1981	Study Type: prospective single cohort Study Quality Score (0–6): 3(111000) OME definition: type B tympanogram signified middle-ear effusion Group: First graders in two municipalities (Hirtshals and Sindal, Denmark) who started school in August, 1978 N=387 children at initial exam with 345 tested at all 10 sessions N=100 with known onset and recovery time. N=31 with unknown onset and recovery time	Time: 8/1978–8/1979 Place: First graders in two rural municipalities (Hirtshals and Sindal, Denmark) [24 classes at 14 different schools] Inclusion: • started first grade in 8/1978 in two rural municipalities (Hirtshals and Sindal, Denmark) Exclusion: None	Type of OME: newly diagnosed OME of unknown duration; OME persisting for weeks or months Age: 6 1/2–7 1/2 year (first graders) Examiner(s): specially trained audiometricians performed the tympanogram and audiology testing OME diagnostic method: tympanometry • A: pressure > –100 mmH_2O • B: otoadmittance < 0.20 millimhos, absolute gradient < 0.04 millimhos, and absence of ipsilateral acoustic reflex • C: ≤ –100 mmH_2O Interventions: The authors did not mention control of medications. One child had adenoidectomy and paracentesis and three had adenoidectomy alone. Interval of screening: 1, 3, 4, 6, 7, 8, 9, 10, 12 months after initial exam	OME resolution rate by **case/child** **<3 months resolution – all subjects** Onset Season # resolved/# at risk (%) Sep–Feb 46/69 (66.7%) Mar–Aug 60/70 (86.0%) **<3 months resolution – on subjects with known onset or resolution time only, excluding unknowns[a]** Onset Season # resolved/# at risk (%) Sep–Feb 46/63 (73.0%) Mar–Aug 41/45 (91.1%) [a] [Onset and resolution time unknown in 3, onset time unknown in 19, and resolution time unknown in 9.]

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
9 Marchisio, Principi, Passali, Salpietro, Boschi, Chetri, Caramia, Longhi, Reali, Meloni, DeSantis, Sacher, and Cupido 1998	Study Type: prospective single cohort embedded in a randomized controlled study Study Quality Score (0–6): 3(111000) OME definition: asymptomatic middle-ear effusion, demonstrated by an abnormal appearance of the tympanic membrane, diffusely opaque, with impaired mobility or presence of air-fluid levels associated with a flat, type B tympanometric curve Group: Children attending the first year of primary school in one of eleven primary schools in Italy N=485 children with OME from 3413 screened. N=62 children with OME at 12-week follow-up randomized to placebo.	Time: October–January during 1993–4 and 1994–5 Place: Eleven primary schools visited by an otoscopist from either one of 9 Pediatric or 2 Otolarynogology departments in different regions of Italy Inclusion: • Attending year 1 in one of eleven participating primary schools in Italy • persistent unilateral or bilateral middle-ear effusion Exclusion: • craniofacial abnormality • any major congenital malformation • serious underlying disease • acute upper respiratory infection including AOM • high risk of sensorineural hearing loss • chronic suppurative OM • perforation of tympanic membrane • previous ear surgery	Type of OME: newly diagnosed OME of unknown duration; OME persisting for weeks or months, i.e. at least 12 weeks Age: 16 were 5 years old, 211 were 6 years old, and 258 were 7 years old Gender: 248 males, 237 females Laterality: 219 unilateral, 262 bilateral Month of exam: 296 in Oct–Nov, 189 in Dec–Jan Examiner(s): validated otoscopist OME diagnostic method: pneumatic otoscopy • abnormal: abnormal tympanic membrane appearance with impaired mobility or air-fluid levels and portable tympanometry • B: flat curve Interventions: It appears none of the children received antibiotics in the last month of the follow-up period and did not receive any surgical intervention during the follow-up interval. Interval of screening: 12 weeks after initial diagnosis; then 16 and 20 weeks after initial diagnosis for children randomized to placebo after 12 weeks	OME resolution by **child** (Page 559) Resolution Time # resolved/# at risk (%) <12 weeks 331/451 (73.4%) <16 weeks[a] complete 7/59 (11.9%) partial 13/59 (22.0%) <20 weeks[a] complete 11/59 (18.6%) partial 17/59 (28.8%) [a] OME resolution (time after initial diagnosis, i.e. 12 weeks plus time after randomization to placebo) of subgroup of children with OME at 12-week follow-up randomized to placebo. (Table III)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
1927 Mills and Vaughan-Jones 1992	Study Type: prospective single cohort embedded in a prospective comparative cohort Study Quality Score (0–6): 1(100000) OME definition: not defined Group: New cases of childhood OME (patients < 15 years of age) who presented at principal author's clinic between 10/1986–10/1988 N=192 children	Time: 10/1986–10/1988 Place: Clinics, Department of Otolaryngology, Ninewells Hospital and Medical School, Dundee, UK Inclusion: • <15 years old Exclusion: • impedance findings suggested effusion was in fact due to thickening of the tympanic membrane	Type of OME: newly diagnosed OME of unknown duration Age: 1–14 years old, mean 5.5 years Examiner(s): unknown OME diagnostic method: pneumatic otoscopy tympanometry Interventions: Most of the children had myringotomies at the 2 month follow-up. Two children had "conservative treatment" at some unspecified point in time. Interval of screening: usually 2 months after initial exam	OME resolution by **child** (Figure 4) Resolution Time # resolved/# at risk (%) <2 months 57/192 (29.7%) [Authors reported uncertainty in the timing of the follow-up visit.]

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
1946 Moller and Tos 1990	Study Type: prospective single cohort Study Quality Score (0–6): 3(111000) OME definition: not defined Group: Healthy children attending a kindergarten in Copenhagen, Denmark N=51 children, 100 ears	Time: 11/1/1989–11/30/1989 (21 days of screening) Place: A kindergarten in Copenhagen, Denmark; Gentofte Hospital/University of Copenhagen Inclusion: • kindergartners Exclusion: • ears with grommets inserted (two ears excluded) • definite type of tympanogram could not be identified • child defected from study or refused the examination w/either one or the other instrument	Type of OME: newly diagnosed OME of unknown duration Age: 37–68 months old Gender: 21 males, 30 females Examiner(s): unknown OME diagnostic method: tympanometry • A: +99 to –99 mmH2O • B: • AZ 7: flat curve without impedance minimum or with a measurable impedance minimum and relative gradient below 0.1 • ZS 331: flat training or compliance below 0.25 ml and absent ipsilateral stapedial reflex • C1: –100 to –199 mmH_2O • C2: >–200 mmH_2O Interventions: The authors did not mention control of interventions. Interval of screening: daily for 30 days	OME resolution by **ear** for OME noted at start of study (page 938 and 939 and Tables IV and V) **AZ7[Impedance audiometer]** Time Resolved Type # resolved/# at risk (%) <3 days B to A 1/4 (25.0%) B to A/C1 1/4 (25.0%) <10 days B to A 1/4 (25.0%) B to A/C1 3/4 (75.0%) <15 days B to A not done B to A/C1 3/4 (75.0%) <30 days B to A 0/2 (0.0%) B to A/C1 1/2 (50.0%) **ZS331 [Impedance tympanoscope]** Time Resolved Type # resolved/# at risk (%) <3 days B to A 2/7 (28.6%) B to A/C1 2/7 (28.6%) <10 days B to A 4/7 (57.1%) B to A/C1 6/7 (85.7%) <15 days B to A not done B to A/C1 5/6 (83.3%) <30 days B to A 0/3 (0.0%) B to A/C1 1/3 (33.3%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
2184 Portoian-Shuhaiber and Cullinan 1984	Study Type: prospective single cohort Study Quality Score (0–6): 3(111000) OME definition: abnormal defined as an abnormal tympanometric curve and/or absent acoustic reflex Group: Children, 5–6 years old, attending selected primary schools in SE London N=130 children with OME from 318 screened	Time: March 1979 and May 1979 Place: Twelve primary schools in SE London Inclusion: • 5–6 years old • attending one of the twelve selected schools in SE London Exclusion: • conditions predisposing to glue ear such as Down's syndrome and cleft palate	Type of OME: newly diagnosed OME of unknown duration Age: 5–6 years Ethnicity: African-American: 57; Indian: 16; White: 234; unknown: 11. Examiner(s): unknown OME diagnostic method: tympanometric and acoustic reflex measurements • abnormal: abnormal curve and/or absent reflex • normal: normal tympanogram curve with positive acoustic reflex Interventions: The authors did not mention control of interventions. Interval of screening: 10 weeks after initial diagnosis	OME resolution by **child** (page 1112) **<10 weeks resolution** Ethnicity # resolved/# at risk (%) Total 65/130 (50.0%) African-American 5/7 (71.4%) Indian 5/9 (55.6%) White 55/110 (50.0%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 2627 and 2634 for further findings]

2189
Poulsen and Tos
1978
[This cohort is also in 2627, 2631, 2634, 2639, 2642, 4834, and 4835.]
	Study Type: prospective single cohort
	
	Study Quality Score (0–6): 3(111000)
	
	OME definition: not defined
	
	Group: Infants born between January–April, 1977, in the Gentofte maternity ward
	
	N=151 screened initially, 109 children (218 ears) at all examinations.
	
	N=0 for type B or C2 at cohort inception.
		Time: recruitment 1/1997 to 4/1997 with last follow-up 6 months after initial exam
		
		Place: Copenhagen, Denmark
		
		Inclusion:
		• healthy newborns born between January and April 1977 in the Gentofte maternity ward
		
		Exclusion: None
			Type of OME: newly diagnosed OME of unknown duration
			
			Age: Newborns
			
			Gender: 82 males, 69 females
			
			Examiner(s): unknown
			
			OME diagnostic method: tympanometry
			• A: >–100 mmH_2O
			• B: flat curve with impedance slope ≤0.1
			• C: –100 to –300 mmH_2O with impedance slope >0.1
			  • C1: –100 to –199 mmH_2O
			  • C2: –200 to –350 mmH_2O
			
			Interventions: The authors did not mention control of interventions.
			
			Interval of screening: 3 and 6 months after initial diagnosis at 2–4 days of age
				OME resolution by **ear** (Figure 1)
				
				Resolution from C1 to A
				
				Time # resolved/# at risk (%)
				<3 months 19/24 (79.2%)
				<6 months 13/24 (54.2%)
				
				Cumulative OME resolution by **ear** (Figure 1)
				
				Resolution from C1 to A
				
				Time # resolved/# at risk (%)
				<3 months 19/24 (79.2%)
				<6 months 21/24 (87.5%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 2629, 2634, and 2593 for further findings]
2190 Poulsen and Tos 1980 [This cohort is also in articles 2629, 2631, 2634, 2593, 2639, 2642, 4834, and 4835.]	Study Type: prospective single cohort Study Quality Score (0–6): 3(111000) OME definition: not defined Group: Healthy two-year-old children born between the first and tenth day of every month in 1976 residing in two municipalities in the northern part of Copenhagen county N=278 children at initial exam but N=240 children who presented at both follow-up exams N for type B at cohort inception=60. N for type B/C2 at cohort inception=172. N for type B/C2/C1 at cohort inception=290.	Time: 11/1977–5/1978 Place: Two municipalities in the northern part of Copenhagen county, Denmark Inclusion: • Age: 2 years • born between the 1st and the 10th day in every month of 1976 in 2 Copenhagen county municipalities • healthy Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: 2 year old at final exam Examiner(s): unknown OME diagnostic method: tympanometry • A: 0 to –99 mmH_2O • B: flat curve • C1: –100 to –199 mmH_2O • C2: –200 to –350 mmH_2O Interventions: The authors did not mention control of medications. Three children received ear tubes, and three had adenoidectomy. Interval of screening: 3 and 6 months after initial diagnosis	OME resolution by **ear** using the initial Nov. 1977 denominators from Table I in article 2631 (Table III) **<6 month resolution** Type # resolved/# at risk (%) B to A 15/59 (25.4%) B to A/C1 25/59 (42.4%) B to A/C1/C2 38/59 (64.4%) B/C2 to A 37/152 (24.3%) B/C2 to A/C1 74/152 (48.7%) B/C2/C1 to A 87/248 (35.1%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
2240 Renvall, Lidén, Jungert, and Nilsson 1978	Study Type: prospective single cohort Study Quality (0–6): 3 (111000) OME definition: not defined. Group: Children aged 10–11 years who were initially examined at age 7 [only ears with a middle ear pressure <=–100 mm H2O or a flat tympanogram at the initial exam were included] N=210 children, 335 ears	Time: not specified Place: Göteborg, Sweden Inclusion: • 10–11 years (initially evaluated at age 7) • only includes ear that at the initial evaluation (age 7) had middle ear pressure of <=–100 mmH_2O or a flat tympanogram Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: 10–11 years (initially evaluated at age 7) Examiner(s): otologic examination by experienced otologist, audiologic examiner unknown OME diagnostic method: tympanometry and stapedius reflex measurements and otoscopic exam initial screen • abnormal: ≤–100 mmH_2O or flat tympanogram follow-up • abnormal: ≤–150 mmH_2O or flat tympanogram, stapedius reflex threshold > 95 dB H.L., and effusion by otologic exam Interventions: The authors did not mention control of interventions. Interval of screening: 3 years after initial diagnosis	OME resolution by **ear** (Table III and I) **<3 years resolution** Diagnostic Method # resolved/# at risk (%) Tympanogram 282/335 (84.2%) Otologic exam 250/335 (74.6%) [The authors concluded that "reduced middle ear pressure is more common in ears once having had subnormal middle ear pressure than in a nonselected group. They did not suggest that these necessarily represent persistent OME.]

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
2242 Renvall, Anniansson, and Lidén 1982	Study Type: prospective single cohort Study Quality (0–6): 4(111100) OME definition: not defined Group 3: 4-year-old children with abnormal tympanograms after hearing screen and comprehensive evaluation N=5928 initially screened and N=223 ears with abnormal tympanogram. N for type B at cohort inception=58 ears. N for type B/C at cohort inception=223 ears.	Time: During 1980 Place: S1~Healthy-baby-clinics in Goteborg, Sweden; S2~children who failed screening were referred to two oto-audiological health centers in Goteborg Inclusion: • Age: 4 years • S2~failure criterion >20db Exclusion: None	Type of OME: Newly diagnosed OME of unknown duration Age: 4 years Examiner(s): unknown OME diagnostic method: tympanometry, audiology, and otologic exams • abnormal: middle-ear pressure ≤–150 mmH_2O, hearing ≤20 dB HL, and/or retraction pocket on otologic exam • normal: middle-ear pressure > –150 mmH_2O, no tympanic membrane retraction, and hearing <= 20 dB HL Tympanogram classification: • Flat to >150 mmH_2O [similar to type B to A] • Flat or –150 to –400 mmH_2O to >=150 mmH_2O [similar to type B or C to A] Interventions: The children did not receive any treatment for four months after initial diagnosis. Interval of screening: 6 and 12 weeks after initial diagnosis. Those with normal exams at any point did not receive follow-up.	Cumulative OME resolution by **ear** (Figures 2 and 3) **<6 weeks resolution** Type # resolved/# at risk (%) B to A 10/40 (25.0%) B/C to A 56/144 (38.9%) **<12 weeks resolution** Type # resolved/# at risk (%) B to A 16/40 (40.0%) B/C to A 86/144 (59.7%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
2243 Reves, Budgett, Miller, Wadsworth, and Haines 1985	Study Type: prospective single cohort Study Quality (0–6): 3(111000) OME definition: not defined Group: Children 3 months to 6 years old on 10/31/1983 recruited from age-sex register and successfully tested by tympanometry N=232 children, 452 ears, of whom n=220 had adequate exams	Time: 11/1983–2/1984 Place: General practice at a health centre in north west London (situated on a council housing estate and serves a deprived population) Inclusion: • 3 months–6 years on 10/31/1983 • listed in age-sex register Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: 3 months–6 years on 10/31/1983; <1 year old 59 ears, 1 year old 73 ears, 2 years old 94 ears, 3 years old 86 ears, 4 years old 82 ears, 5 years or older 58 ears Laterality: unilateral 29, bilateral 39 Examiner(s): unknown OME diagnostic method: tympanometry • A: normal pressure (–100 to +50 mmH_20), normal compliance (>0.3), peak at normal pressure • As: normal pressure, low compliance, small peak at normal pressure • B: low compliance and middle-ear pressure < –100 mmH_20 or indeterminate, usually flat or peak < –100 mmH_20 (signified middle-ear effusion) • C: negative pressure, normal compliance, peak at negative pressure Interventions: The authors did not mention control of interventions. Interval of screening: every 2–4 weeks after initial diagnosis for 6 months	OME resolution by **child** (comment in abstract) Resolution Time # resolved/# at risk (%) <3 months 40/68 (58.8%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
2262 Roberts, Johnson, Carlin, Turczyk, Karnuta, and Yaffee 1995	Study Type: prospective single cohort Study Quality Score (0–6): 4(111010) OME definition: not defined Group: Healthy, full-term infants born in MetroHealth Medical Center, Cleveland, Ohio, were enrolled on day 1 of life N=68 subjects	Time: not specified Place: Normal newborn nursery of MetroHealth Medical Center, a county hospital in Cleveland, Ohio, and a county pediatric clinic Inclusion: • healthy, full-term newborn Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: Newborns (max follow-up ~ 2months) Race/ethnicity: white 47%, African American 43%, Hispanic 10% Examiner(s): 2 otoscopists; tympanometry and acoustic reflex measurements by an audiologist on day 1 and 3 and by a research assistant at 2 weeks and 2 months OME diagnostic method: pneumatic otoscopy • middle-ear effusion: tympanic membrane mobility markedly decreased or air-fluid level tympanometry • middle-ear effusion: peak susceptance < 0 millisiemens • no effusion: > 0 millisiemens acoustic reflex measurement • middle-ear effusion: threshold at 110dB HL or absent reflex • no effusion: threshold up to 100dB HL Interventions: Asymptomatic infants were not treated with antibiotics. No indication of how many infants required intervention. Interval of screening: 2 weeks and 2 months after birth of infants whose parents chose to receive well-child care from the research team	Cumulative OME resolution by otoscopy by **ears** (page 875) Resolution Time # resolved/# at risk (%) <2 weeks 22/24 (91.7%) <2 months 24/24 (100%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
2270 Robinson, Allen, and Root 1988	Study Type: prospective single cohort Study Quality (0–6): 1(110000) OME definition: not defined Group: Children, primarily lower socioeconomic status infants between 3–6 months of age, who underwent tympanometry as part of the study N=137 subjects initially tested of whom n=63 had OME). N for B/C/Cs at cohort inception=63.	Time: not specified Place: 12 Health Department clinics in two metropolitan Detroit counties (Wayne and Oakland); (8 Oakland County clinics are "well baby" clinics; 4 Wayne County clinics serve Medicaid patients) Inclusion: • Age: 6–13 months • income limits are imposed to qualify for Tx at all clinics involved Exclusion: None	Type of OME: Age: 6–13 months, mean 9.7 months (standard deviation 2.3 months) Gender: 65 male, 72 female Race/ethnicity: 63 black, 74 white Examiner(s): second-year audiology master of arts student OME diagnostic method: tympanometry • A: –149to +50 mmH_20 and compliance > 0.2 ml • failures: all other types As, C, Cs, and B Interventions: The authors did not mention control of interventions. Interval of screening: minimum 6 weeks from initial screen if failed	OME resolution by **ear** (pages 343–344) **<6 weeks resolution**[a] Type # resolved/# at risk (%) B to A 10/25 (40.0%) B to A/As 11/25 (44.0%) B to A/As/Cs 13/25 (52.0%) B/C to A 22/41 (53.7%) B/C to A/As 23/41 (56.1%) B/C to A/As/Cs 24/41 (58.5%) B/C/Cs to A 20/34 (58.8%) [a]-minimum 6-week interval but exact intervals not reported

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
2457 Sly, Zambie, Fernandes, and Frazer 1980	Study Type: two prospective single cohorts Study Quality Score (0–6): 3(111000) OME definition: not defined Group 1: February, 1977 screening children in 4–5 year old classes at Memorial Baptist Kindergarten (n=94 children, 188 ears) Group 2: September, 1977 screening children in 4–5 year old classes at Memorial Baptist Kindergarten N1=94 children, 188 ears N2=94 children, 188 ears	Time: February and September, 1977 Place: Memorial Baptist Kindergarten, New Orleans, LA; from the Dept. of Pediatrics, Louisiana State University Medical Center Inclusion: • Age: 4–5 years • written parental permission & child willingness Exclusion: • presence of functional ventilating tube (ear) • not able to obtain adequate seal in the ear canal (ear)	Type of OME: newly diagnosed OME of unknown duration Age: Kindergartners (4–5 years) Examiner(s): unknown OME diagnostic method: tympanometry • abnormal: flat type B without discernible peak in compliance with decrease in pressure to –400 mmH_2O, compliance < 0.3 cc, or peak compliance occurred at or below pressure –100 mmH_2O Interventions: Group 1 antibiotics with or without antihistamines and decongestants in 2 children and antihistamines and decongestants in 9, Group 2 antibiotics with or without antihistamines and decongestants in 5 children and antihistamine and/or decongestants in 21. Interval of screening: every two weeks after initial diagnosis for six weeks	OME resolution by **child/ear** as unit of measure (only those without any treatment). (Table VII) **<2 weeks resolution – February 1997** Unit Type # resolved/# at risk (%) Child B to A 1/7 (14.3%) Child B/C to A 8/24 (33.3%) Ear B to A 1/9 (11.1%) Ear B/C to A 8/32 (25.0%) **<2 weeks resolution – September 1997** Unit Type # resolved/# at risk (%) Child B to A 0/3 (0.0%) Child B/C to A 2/14 (14.3%) Ear B to A 0/5 (0.0%) Ear B/C to A 4/22 (18.2%) **<4 weeks resolution – February 1997** Unit Type # resolved/# at risk (%) Child B to A 3/7 (42.9%) Child B/C to A 12/24 (50.0%) Ear B to A 4/9 (44.4%) Ear B/C to A 14/32 (43.8%) **<4 weeks resolution – September 1997** Unit Type # resolved/# at risk (%) Child B to A 0/3 (0.0%) Child B/C to A 3/14 (21.4%) Ear B to A 0/5 (0.0%) Ear B/C to A 5/22 (22.7%) *Continued on next page*

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
Continued 2457 Sly, Zambie, Fernandes, and Frazer 1980				*Continued* **<6 weeks resolution – February 1997** Unit Type # resolved/# at risk (%) Child B to A 5/7 (71.4%) Child B/C to A 16/24 (66.7%) Ear B to A 6/9 (66.7%) Ear B/C to A 18/32 (56.3%) **<6 weeks resolution – September 1997** Unit Type # resolved/# at risk (%) Child B to A 0/3 (0.0%) Child B/C to A 7/14 (50.0%) Ear B to A 0/5 (0.0%) Ear B/C to A 11/22 (50.0%) Note: Denominator at each time is most likely same as total with OME, however defined.

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 2629, 2190, and 2634 for further findings]
2593 Thomsen and Tos 1981 [This cohort is also in articles 2190, 2629, 2631, 2634, 2639, 2642, 4834, and 4835.]	Study Type: [a] Study Quality (0–6): [a] OME definition: [a] Group: [a] N=184 children who presented at all 5 follow-up exams N for type B at cohort inception=60. N for type B/C2 at cohort inception=172. N for type B/C2/C1 at cohort inception=477.	Time: 11/1977 to 2/1980 Place: [a] Inclusion: [a] Exclusion: [a]	Type of OME: [a] Age: [a] Gender: males 92, females 92 Examiner(s): unknown OME diagnostic method: [a] Interventions: [a] Interval of screening: extended to 5 years old	OME resolution by **ear** (Table IV and Figure 1) **<2 years resolution** Type — # resolved/# at risk — (%) B to A — 9/48 — (18.8%) B to A/C1 — 19/48 — (39.6%) B to A/C1/C2 — 40/48 — (83.3%) B/C2 to A — 26/121 — (21.5%) B/C2 to A/C1 — 53/121 — (43.8%) B/C2/C1 to A — 50/199 — (25.1%) **Resolution based on B to A/C1/C2** Resolution Time — # resolved/# at risk — (%) <3 months — 25/48 — (52.1%) <6 months — 27/48 — (56.3%) <9 months — 39/48 — (81.3%) <12 months — 35/48 — (72.9%) <24 months — 40/48 — (83.3%) Cumulative OME resolution by **ear** **Resolution based on B to A/C1/C2** Resolution Time — # resolved/# at risk — (%) <3 months — 25/48 — (52.1%) <6 months — 30/48 — (62.5%) <9 months — 39/48 — (81.3%) <12 months — 42/48 — (87.5%) <24 months — 47/48 — (97.9%)

[a] See 2190 Poulsen and Tos (1980).

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 2189 and 2634 for further findings]
2627 Tos 1979 [This cohort is also in articles 2189, 2631, 2634, 2639, 2642, 4834, and 4835]	Study Type: [a] Study Quality Score (0–6): [a] OME definition: [a] Group: [a] N=90 children who presented for all 5 follow-up exams N for type B at cohort inception=4 in this article but was 3 in articles 3189 and 2631. N for type B/C2 at cohort inception=27. N for type B/.C2/C1 at cohort inception=91.	Time: recruitmen1t 1/1997 to 4/1997 with last follow-up 12 months after initial exam Place: [a] Inclusion: [a] Exclusion: [a]	Type of OME: In this evidence table, we are looking at those from the original cohort with tympanogram type B (n=4 in article 2627 and n=3 in article 2189), C1 (n=20 in article 2189), or C2 (n=58 in article 2189) at the 6-month follow-up visit and presented for all 5 follow-up visit. Type B was not seen until the 6-month follow-up visit. Type C1 was seen in 24 patients at 2–4 days of age and 32 at the 3-month follow-up. Type C2 was seen initially in 1 patient at the 3-month follow-up. [a] Age: [a] Gender: 50 males, 40 females Examiner(s): [a] OME diagnostic method: [a] Interventions: [a] Interval of screening: 1, 3, 6, 9, and 12 months after initial diagnosis at 2–4 days of age. [a]	OME resolution by **ear** using the 6-month exam denominators from Table IV in article 2189 (Tables III and V) **<3 months resolution** Type # resolved/# at risk (%) B to A 1/4 (25.0%) B to A/C1 1/4 (25.0%) B to A/C1/C2 1/4 (25.0%) B/C2 to A 5/20 (25.0%) B/C2 to A/C1 11/20 (55.0%) B/C2/C1 to A 24/75 (32.0%) **<6 months resolution** Type # resolved/# at risk (%) B to A 1/4 (25.0%) B to A/C1 1/4 (25.0%) B to A/C1/C2 1/4 (25.0%) B/C2 to A 2/20 (10.0%) B/C2 to A/C1 8/20 (40.0%) B/C2/C1 to A 25/75 (33.3%)

[a] See 2189 Poulsen and Tos (1978).

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 2190, 2634, and 2593 for further findings]
2629 Tos 1979 [This cohort is also in articles 2190, 2631, 2634, 2593, 2639, 2642, 4834, and 4835.]	Study Type: [a] Study Quality Score (0–6): [a] OME definition: [a] Group: Healthy two-year-old children residing in two Copenhagen county municipalities N=278 children at initial exam; N=222 children who presented at all 3 follow-up exams. N for type B at cohort inception=60. N for type B/C2 at cohort inception=172. N for type B/C2/C1 at cohort inception=280.	Time: 11/1977–8/1978 [a] Place: [a] Inclusion: [a] Exclusion: [a]	Type of OME: [a] Age: [a] Examiner(s): [a] OME diagnostic method: [a] Interventions: [a] Interval of screening: 3, 6, and 9 months after initial diagnosis [a]	OME resolution by **ea**r using the initial 11/1977 denominators from Table I in article 2631 (Table III) **<9 months resolution** Type # resolved/# at risk (%) B to A 19/51 (37.3%) B to A/C1 36/51 (70.6%) B to A/C1/C2 43/51 (84.3%) B/C2 to A 63/138 (45.7%) B/C2 to A/C1 100/138 (72.5%) B/C2/C1 to A 118/229 (51.5%)

[a] See 2190 Poulsen and Tos (1980).

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 2629, 2190, and 2593 and 2189 and 2627 for further findings]
2634 Tos 1980 [The 1976 cohort is also in articles 2190, 2629, 2631, 2593, 2639, 2642, 4834, and 4835. The 1977 cohort is also in articles 2189, 2627, 2631, 2639, 2642, 4834, and 4835.]	Study Type: [a] Study Quality Score (0–6): [a] OME definition: [a] Group: [a] N=[a]	Time: [a] Place: [a] Inclusion: [a] Exclusion: [a]	Type of OME: [a] Age: [a] Examiner(s): [a] OME diagnostic method: [a] Interventions: [a] Interval of screening: [a]	OME resolution by **ear** (Table III) **<3 months resolution** Type # resolved/# at risk (%) B to A 6/51 (11.8%) B to A/C1 13/51 (25.5%) B to A/C1/C2 27/51 (52.9%) **<6 months resolution** Type # resolved/# at risk (%) B to A 15/51 (29.4%) B to A/C1 21/51 (41.2%) B to A/C1/C2 30/51 (58.8%) **<9 months resolution** Type # resolved/# at risk (%) B to A 19/51 (37.3%) B to A/C1 36/51 (70.6%) B to A/C1/C2 43/51 (84.3%) Cumulative OME resolution by **ear** **<3 months resolution** (same as above) **<6 months resolution** Type # resolved/# at risk (%) B to A 16/51 (31.4%) B to A/C1 24/51 (47.1%) B to A/C1/C2 34/51 (66.7%) **<9 months resolution** Type # resolved/# at risk (%) B to A 24/51 (47.1%) B to A/C1 34/51 (66.7%) B to A/C1/C2 44/51 (86.3%)

[a] See articles 2190 Poulsen and Tos (1980) and 2189 Poulsen and Tos (1978).

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 1486 for further findings]
543 Tos, Holm-Jensen, Sørenson, and Mogensen 1982 [This cohort is also in articles 1486, 2636, 2639, 2642, 4834, and 4835.]	Study Type: [a] Study Quality (0–6): [a] OME definition: [a] Group: [a] N=288 children who attended all 4 follow-up exams N for type B at cohort inception=101 ears (note difference from article 1486). N for type B/C2 at cohort inception=317 ears (note difference from article 1486). N for type B/C2/C1 at cohort inception=477 ears.	Time: The authors identify this study as their 1976 birth cohort, but it is actually their 1975 birth cohort. [a] Place: [a] Inclusion: [a] Exclusion: [a]	Type of OME: [a] Age: [a] Examiner(s): two experienced technicians [a] OME diagnostic method: [a] Interventions: [a] Interval of screening: [a]	OME resolution by **ear** **<3 months resolution** (Figure 1) Type # resolved/# at risk (%) B to A 3/87 (3.4%) B to A/C1 12/87 (13.8%) B to A/C1/C2 51/87 (58.6%) B/C2 to A 49/265 (18.5%) B/C2 to A/C1 102/265 (38.5%) B/C2/C1 to A 103/393 (26.2%) **<6 months resolution** (Figure 1) Type # resolved/# at risk (%) B to A 14/87 (16.1%) B to A/C1 22/87 (25.3%) B to A/C1/C2 49/87 (56.3%) B/C2 to A 51/265 (19.2%) B/C2 to A/C1 118/265 (44.5%) B/C2/C1 to A 96/393 (24.4%) **<9 months resolution** (Figures 3 and 4) Type # resolved/# at risk (%) B to A/C1 20/82 (24.4%) B to A/C1/C2 39/82 (47.6%) B/C2 to A/C1 106/247 (42.9%) **<12 months resolution** (Figures 3 and 4) Type # resolved/# at risk (%) B to A/C1 15/82 (18.3%) B to A/C1/C2 42/82 (51.2%) B/C2 to A/C1 76/247 (30.8%)

[a] See 1486 Holm-Jensen, Sørenson, and Tos (1981) and 2636 Tos (1981).

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings [See 1486 for further findings]
543 Tos, Holm-Jensen, Sørenson, and Mogensen 1982 [This cohort is also in articles 1486, 2636, 2639, 2642, 4834, and 4835.]				Cumulative OME resolution by **ear** **<3 months resolution** (same as above) **<6 months resolution** (Figure 1) Type # resolved/# at risk (%) B to A 17/87 (19.5%) B to A/C1 31/87 (35.6%) B to A/C1/C2 62/87 (71.3%) B/C2 to A 85/265 (32.1%) B/C2 to A/C1 152/265 (57.4%) B/C2/C1 to A 152/393 (38.7%) **<9 months resolution** (Figures 3 and 4) Type # resolved/# at risk (%) B to A/C1 37/82 (45.1%) B to A/C1/C2 60/82 (73.2%) B/C2 to A/C1 166/247 (67.2%) **<12 months resolution** (Figures 3 and 4) Type # resolved/# at risk (%) B to A/C1 41/82 (50.0%) B to A/C1/C2 64/82 (78.1%) B/C2 to A/C1 174/247 (70.5%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
91 van Balen, de Melker, Touw-Otten 1996	Study Type: prospective single cohort embedded in a randomized controlled trial Study Quality (0–6): 3(111000) OME definition: presence of fluid in the middle-ear cavity behind an intact tympanic membrane without signs or symptoms of acute infection Group: children with bilateral OME N=433	Time: recruited 12/1992 to 8/1994 and then followed for 3 months Place: 57 general practices in the Netherlands Inclusion: • 6 months to 6 years old • bilateral OME Exclusion: none	Type of OME: newly diagnosed OME of unknown duration Age: 6 months to 6 years old Laterality: all bilateral Examiner(s): general practitioner trained in tympanometry OME diagnostic method: tympanometry • A: –99 to +200 dPa and max compliance >= 0.2 mmho • B: <= –400 dPa and max compliance < 0.2 mmho • C1: –199 to –100 dPa and max compliance >=0.2 mmho • C2: –399 to –200 dPa and max compliance >= 0.2 mmho • Middle-ear effusion defined by B or C2 curves. Interventions: Authors stated children received no treatment for their ear problems and noted that “In the Netherlands it is not common to treat OME with antimicrobial drugs.” It is possible antibiotics were used during the watchful waiting period. Interval of screening: 3 months after initial exam	OME resolution by **child** (Figure) Time # resolved/# at risk (%) **<3 months** 223/433 (51.5%) [Authors did not distinguish between partial and complete resolution.]

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
2791 Williamson 1994	Study Type: prospective single cohort Study Quality Score (0–6): 2(110000) OME definition: not defined Group: Children attending four adjacent Infant and First schools in South West Hampshire participating schools and screened by tympanometry once per term at a four-month interval until they left primary school N=856 subjects	Time: Recruitment 9/1988–3/1989; study period ended summer 1991 Place: School medical room in four adjacent Infant and First Schools in the SW Hampshire area Inclusion: • 5–8 years old • attending four attending four adjacent Infant and First schools in South West Hampshire Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: 5–8 years old Examiner(s): unknown OME diagnostic method: tympanometry • A: 200 to –99 pressure • B: lack of defined compliance maximum • C1: –100 to –199 pressure • C2: –200 to –400 pressure Interventions: The authors report that 27 children had grommets in situ during the study period. This includes 6 children in the Spring 1990 cohort with type B tympanograms whose OME resolution is reported. Interval of screening: once per term at a four-month interval until the child left primary school	OME resolution of a cohort with **type B tympanogram** in Spring of 1990 (We assume the denominator at each time is the total.) (page 932) **Resolution by CHILD** Resolution Time — # resolved/# at risk — (%) <4 months 22/50 (44.0%) <8 months 38/50 (76.0%) <1 year 45/50 (90.0%) **Resolution by EAR** Resolution Time — # resolved/# at risk — (%) <4 months 35/67 (52.2%) <8 months 52/67 (77.6%) <1 year 61/67 (91.0%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Inclusion/Exclusion Criteria	Influencing Factors	Findings
2863 Zielhuis, Rach, and van den Broek 1990	Study Type: prospective single cohort Study Quality Score (0–6): 3(111000) OME definition: type B tympanogram signified OME but otherwise not defined Group: Preschool children seen in follow-up for OME every 3 months from 2–4 years of age (n=609 children had all nine total screenings out of 1328 whose parents agreed to participate) N=1328 subjects	Time: Follow-up from age 2–4 years in children born between 9/1/1982–8/31/1983 Place: Measurements carried out at the children's home in Nijmegen, The Netherlands Inclusion: • Age: 4 years • born in Nijmegen during the period 9/1/1982–8/31/1983 and living in Nijmegen on their 2nd birthday • screened every 3 months since 2 years of age • parental consent Exclusion: None	Type of OME: newly diagnosed OME of unknown duration Age: 4 years old Examiner(s): three trained audiologic assistants OME diagnostic method: tympanometry • A: compliance ≥0.2 ml and pressure ≥–99 dPa • B: compliance <0.2 ml and pressure ≤ –400 dPa • C1: compliance ≥0.2 ml and pressure –100 to –199 dPa • C2: compliance ≥0.2 ml and pressure –200 to –399 dPa • B: compliance <0.2 ml and pressure ≤ –400 dPa Interventions: The authors did not mention control of medications. 56 of the children were treated with ventilating tubes. Interval of screening: every 3 months after initial screen until 4 years old	The authors presented an equation to estimate resolution of first OME episode by **ear** based on the study data. (page 217) y=2exp(–0.33x) y=probability of OME x=follow-up in months r^2=0.98 Duration of all OME by ear (Figure 4) Mean: 5.0 months Standard error of mean: 0.08 months 5th %ile: 3 months 25th %ile: 3 months median: 3 months 75th %ile: 6 months 95th %ile: 12 months [mixture of OME diagnosed at initial screen and OME arising during the study period] Estimated cumulative OME resolution by **ear** based on the above equation and numbers given in Figure 2 (N=1631 ears of 816 children) Time # resolved/# at risk (%) <3 months 816/1631 (50.0%) <6 months 1223/1631 (75.0%)

Evidence Table 1. Natural History of Otitis Media with Effusion (Continued)

Record# Author Year	Findings
2863 Zielhuis, Rach, and van den Broek 1990	Duration of all OME episodes by ear by influencing factors (Table 3)

Factor	Category	% with duration of 3 months	% with duration of 6 months	% with duration of 9 months or more	Mean Duration	p-value
Season at start of OME	Autumn	59.7%	19.5%	20.8%	4.8	<0.001
	Winter	65.6%	17.7%	16.7%	4.5	
	Spring	61.3%	24.2%	14.5%	4.6	
	Summer	69.0%	20.2%	10.8%	4.3	
Age at end of episode	24-29	77.9%	22.1%	0.0%	3.7	<0.001
	30-35	60.4%	19.5%	20.1%	4.8	
	36-41	64.4%	20.7%	14.9%	4.5	
	42-48	56.6%	19.4%	20.0%	5.0	
Sex	Male	64.4%	20.9%	14.7%	4.5	0.44
	Female	63.7%	19.7%	16.6%	4.6	
Upper respiratory tract infection	Yes	64.4%	19.5%	16.1%	4.6	0.24
	No	63.8%	22.2%	14.0%	4.5	
History of acute otitis media	Yes	62.1%	23.0%	14.9%	4.6	0.77
	No	57.1%	24.5%	18.4%	4.8	

[mixture of OME diagnosed at initial screen and OME arising during the study period]

[Actual counts were not presented and could not be derived.]

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development

[The eight components of study quality score are: study cohort clearly defined; subjects assembled at a uniform time point; pathway of subject entry clearly described; complete follow-up achieved; withdrawals/drop-outs described; objective outcomes used; outcome assessment blinded; and extraneous factors adjusted. 1 indicates presence and 0 indicates absence.]

Record# Author Year	**Study Quality OME Definition Group(s) and Sample Size**	**Time/Place Subject Characteristics Inclusion/Exclusion Criteria**	**Risk Groups Predictors Outcome Measures**	**Findings**
877 Black 1993	Prospective Cohorts Study Study Quality Score (0–8): 3 (00001101) OM Diagnosis: By otologic examination by unspecified examiner. OM Groups: Group 1: OM+: High Infant Otitis Media (at least two cases of otitis media within the first year of life documented by otologic examination) Group 2: OM–: No Infant Otitis Media (Had not experienced otitis media during first year of life) N=31 subjects N1=21 N2=10	Time: 4-year follow-up (actual time of study not specified) Place: Inner-city sample in Baltimore, Maryland Characteristics: • 87% had African-American primary caregivers • 61% had not completed high school • 81% single women • 91% received medical assistance • all children were enrolled in kindergarten or Head Start I Inclusion: • Age: 4 years follow-up of the original cohort of infants • partipated in authors earlier study • Born at term with appropriate weight for gestational age • No medical problems beyond OM Exclusion: None	Age: OM history: first year of life Outcomes: 4–6 years of age OM History: At least 2 episodes of OM within the first year documented of otologic exam. A child could receive credit for only one bout of OM within each 29-day period. Outcome Measures: *Cognitive development*: McCarthy Scales of Children's Abilities (verbal, perceptual processing, quantitative, memory, and motor). Each has a mean score of 50 and a standard deviation (SD) of 10. Scores on the first 3 scales are summed to give a GCI with a mean (SD) of 100 (15). *Language development*: Peabody Picture Vocabulary Test-Revised, an assessment of receptive language. Has mean (SD) of 100 (15).	OM+ OM– Test Mean (SD) Mean (SD) McCarthy Scales Verbal 46.7 (11.5) 41.0 (10.7) Perceptual 47.5 (12.8) 43.2 (11.4) Quantitative 43.6 (8.6) 39.1 (10.5) Memory 46.0 (10.6) 41.5 (12.0) Motor 53.0 (12.4) 36.4 (10.2) GCI[a] 92.6 (19.5) 84.9 (19.3) PPVT–R[b] 82.9 (17.3) 72.4 (17.6) [a] General cognitive index [b] Peabody Picture Vocabulary Test–Revised

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1219 Feagans 1987	Prospective cohort study Study Quality Score (0–8): 5 (11100101) OM Diagnosis: OM diagnosed by 2 pediatricians and 2 nurse practitioners using pneumatic otoscopy. Beginning 1978 tympanometry was used to corroborate the diagnosis Group: Children who were part of a medical and day care intervention project N=44 subjects	Time: 1972–1982 Place: Children were sampled from a day care center in Pennsylvania Characteristics: • Black • Low socioeconomic status; • Attended day care center 50 weeks per year from 6 weeks to 5 years of age. Inclusion: • Age: 5–7 years (measures at both time points) • children who were followed from birth to 2nd grade • biologically normal at birth • high risk for general developmental delays in language and intelligence due to low SES Exclusion: None	Age: OM History: 0–3 year of life Outcomes: at 5–7 years of age OM History: Frequency of OM Duration of OM Outcome Measures: Mean length of utterance (MLU) at age 5 Mean length of utterance (MLU) at age 6 Paraphrase score at age 5 Paraphrase score at age 6 Predictors in the study: Mother's IQ at birth Mother's education Home environment (HOME) at 18 months Language (WPPSI IQ) at age 5 Frequency of OM in 0–3 years of life Duration of OM in 0–3 years of life	Mean (SD) Length of utterance At age 5 10.3 (2.9) At age 7 12.4 (3.8) Paraphrase score At age 5 7.2 (1.4) At age 7 8.0 (1.0) **Multiple regression analysis-Age 5** MLU: R^2 = 0.07, p>0.10, no variable significant. Paraphrase score: R^2 = 0.47, p<0.001, WPPSI IQ most significant followed by frequency of OM. **Multiple regression analysis-Age 7** MLU: R^2 = 0.13, p<0.20; IQ was the only significant predictor (R^2= 0.12, p<0.02). Paraphrase score: R^2 = 0.22, p>0.11; WPPSI and duration of OM were significant predictors (R^2= 0.18, p<0.02)

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1255 Fischler 1985	Retrospective-prospective study Study Quality Score (0–8): 5 (11001110) OM Diagnosis: By medical record review of documented physician's clinical diagnosis. OM Groups: Group 1: HR/REC (HR-High risk status,REC-recurrences of OM attacks after age 2) ~ >=2 attacks by age 2 yr and >=3 attacks after age 2 yr Group 2: HR/NREC ~ >=2 attacks by age 2 yr and <3 attacks after age 2 yr Group 3: NHR/NREC ~ <2 attacks by age 2 yr and <3 attacks after age 2 yr N=167 N1=33 N2=63 N3=71	Time: Start date~7/1974 Place: Four Indian reservations in Arizona Characteristics: • 50% had family income under $5000 per year • 37% had housing below average • 13% primarily Apache speaking at home Inclusion: • Age: 6–8 years • healthy Apache Indian children who had been followed since birth Exclusion: • moved during study period • absent or not tested because of time constraints • medical reasons (specific conditions not specified)	Age: OM history: 0–2 years of life Outcome: at 6–8 years Examiner(s): • Unspecified physicians for diagnosis of otitis media; • School nurse for hearing tests; • Pediatrician/otolaryngologist for otoscopic examinations; • Certified speech pathologists for language testing. OM history: Groups defined by number of OM episodes by age 2 Outcome measures: • Four subtests of the Test of Language Development (TOLD). • Two receptive language subtests (picture vocabulary [PV] and grammatic understanding [GU]) • Two expressive subtests (grammatic completion [GC] and oral vocabulary [OV]) • Articulation: informally scored as normal or suspect • Nonverbal intelligence by the block-design subtest of the Wechsler Intelligence Scale for Children-Revised (WISC-R) [For all subtests raw scores were converted into age standardized scores. Standard scores were based on national norms having mean (SD) of 10 (3).]	Mean (SD) Group 1 Group 2 Group 3 OM risk High Moderate Low TOLD subtest Receptive language PV 5.6 (2.3) 6.8 (2.8) 6.1 (2.7) GU 7.2 (2.2) 8.1 (2.3) 7.9 (2.1) TOLD subtest Expressive language GC 5.0 (1.7) 5.7 (1.9) 5.4 (2.4) OV 6.1 (3.0) 7.7 (2.8) 7.1 (2.7) WISC-R block design Non-verbal intelligence 10.6 (2.9) 10.4 (3.1) 10.0 (2.7)

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1277 Freeark 1992	Retrospective-Prospective Cohort Study Study Quality Score (0–8): 4 (10000111) OM Diagnosis: By whom, not specified; How diagnosed, not specified. (OM history obtained from medical records). OM Groups: Group 1: High OM: Above median of OM severity defined by frequency of OM episodes and days with effusion. Group 2: Low OM: Below median of OM severity defined by frequency of OM episodes and days with effusion. N=54 subjects N1=26 N2=28	Time: not specified Place: university-based pediatric clinic, Michigan Characteristics: • 86% white • 82% of mothers and 91% of fathers with at least some college education • 77% of fathers in lower middle to upper middle class occupations or were full-time graduate students • 75% of mothers employed Inclusion: • Age: 3–4 years • in patient roster of a university-based pediatric clinic • both parents living in home • parents had at least a high school ed. and not receiving public assistance • family had no more than four children • child did not suffer from a chronic illness or disability Exclusion: None	Age: OM history: 0–3 years of life Outcome: 3–4 years Examiner(s): Trained examiners for language testing. OM History: OM severity defined by a) number of separate episodes of otitis and b) total number of days of effusion over the first 3 years. Outcome measures: Verbal Scale Index (VSI) of the McCarthy Scale of Children's Abilities – a summary measure of the child's verbal abilities derived from 6 subtests assessing verbal fluency and memory, word knowledge, and verbal reasoning. Stratification Factor: Parent Verbal Stimulation (PVS) – total frequency of mothers' and fathers' descriptive and reflective statements and questions during the first 10 minutes of the dyadic interaction.	VSI below mean (# and %) High PVS Group: High OM 2/14 (14%); Low OM 10/14 (71%) Low PVS Group: High OM 8/12 (67%); Low OM 4/14 (29%) Total: High OM 10/26 (38%); Low OM 14/28 (50%) Hierarchical multiple regression analysis showed that interactions between parent verbal stimulation and otitis severity were significant in predicting the VSI and the VIP score for the interaction segment with fathers.

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1373 Gravel 1992 Same cohort in 1941, 2295 and 4728. See 1941, 2295, and 4728 for further comments.	Prospective Cohort Study Study Quality Score (0–8): 5 (11001110) OM Diagnosis: Pediatric nurse practitioners completed pneumo-otoscopic examinations. OM Groups: Group 1: OM+: Otitis positive children – when bilateral OM was detected at 30% or more of the baby's first year visits. (Also, had poorer auditory sensitivity by click ABR.) Group 2: OM–: Otitis negative children – when middle ear status was rated as normal in both ears during 80% or more of the first year visits. N=23 N1=10 N2=13	Time: not specified Place: Clinical Research Center for Communicative Disorders (which draws subjects from the LIFE Program of the R.F. Kennedy Center, Albert Einstein College of Med. Bronx, NY Characteristics: • low socioeconomic urban neighborhoods based on Hollingshead's (1975) index • none had sensorineural hearing loss Inclusion: • Age: 4 years (measures taken during first year of life as well) • all children were enrolled at 40 wks postconceptional age for follow-up through the LIFE Program • either high-risk infant graduates of a NICU or FT graduates of the well-baby nursery of Jacobi Hospital • English reported as the primary language spoken Exclusion: • neurologically compromised	Age: OM history: 0–1 of life Outcome: at 4 years of age Examiner(s): Not specified OM history: OM groups were defined by otoscopic histories. Outcome measures: *Cognitive measures* (Stanford-Binet 4th Edition): Composite IQ Score Verbal reasoning Abstract/visual reasoning Quantitative reasoning, Short-term memory *Language measure* (Sequenced Inventory of Communication Development-Revised (SICD-R) – a standardized measure of communicative functioning combining parental report with direct assessment: Expressive scale Receptive scale	Mean (SD) OM+ OM– Cognitive N 9 13 Global IQ 87.8 (14.8) 86.0 (8.6) Verbal IQ 88.3 (15.9) 84.3 (9.4) Expressive language in months N 8 12 Mean (SD) 36.0 (5.2) 39.0 (6.2) Receptive language in months N 8 13 Mean (SD) 35.5 (5.4) 37.8 (5.3)

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
4728 Gravel 1996 Same study subjects as 1373, 1941, 2295 See 1373, 1941, and 2295 for further comments.	Prospective Cohort Study Study Quality Score (0–8): 3 (01000110) OM Diagnosis: Pediatric nurse practitioners completed pneumo-otoscopic examinations. OM Groups: Group 1: Otitis positive children – when bilateral OM was detected at 30% or more of the baby's first year visits. Group 2: Otitis negative children – when middle ear status was rated as normal in both ears during 80% or more of the first year visits. N=17 subjects N1=10 N2=7	Time: not specified Place: Clinical Research Center for Communicative Disorders (which draws subjects from the LIFE Program of the R.F. Kennedy Center, Albert Einstein College of Med. Bronx, NY Characteristics: • low socioeconomic urban neighborhoods based on Hollingshead's (1975) index • all had normal hearing and normal middle-ear function on test day Inclusion: • Age: 9 years • born in the same urban hospital • auditory testing conducted during first year of life and at 9 years of age Exclusion: None	Age: OM history: 0–1 of life Outcome: at 9 years of age Examiner(s): Not specified OM history: OM groups were defined by otoscopic histories. Outcome measures: • Binaural masking level difference (MLD) • Speech-in-competition task-adapted from the Pediatric Speech Intelligibility (PSI) test • Selected auditory attention task – An experimental version of the PSI • Story memory tasks – a subset of the standardized Wide Range Assessment of Memory and Learning (WRAML) • Clinical Evaluation of Language Fundamentals-Revised (CELF-R)	No raw data were reported. The following were reported between OM+ and OM– groups: 1. A significant difference between the early OM+ and early OM– groups was found on the story-recall memory task (t=2.42; p=0.032). 2. No significant differences between the groups were found on the MLD, competitive listening tasks, language screen, the PSI selective auditory attention task, or the story recognition-memory task.

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1435 Harsten 1993	Prospective Cohort Study Study Quality Score (0–8): 6 (11101110) OM Diagnosis: AOM was diagnosed by otomicropscopy, performed by an otolaryngologist and defined as an acute episode of earache in a child, with red bulging eardrum(s) or purulent discharge, occasionally febrile and with signs of upper respiratory tract infection. OM Groups: Group 1: Children with recurrent AOM (RAOM defined as at least six episodes of AOM during a 12-month period) during the first 3 years of life Group 2: Children without any AOM episode during the first 3 years of life N=42 subjects N1=13 N2=29	Time: not specified Place: University Hospital of Lund, Sweden Characteristics: • 45% first-borns • 71% mother educated at college level Inclusion: • Age: Birth cohort • monolingual, Swedish children, born at University Hospital of Lund Exclusion: None	Age: OM history: 0–3 year of life Outcomes: at 4 and 7 year of age Examiner(s): Otolaryngologist for otomicroscopy, tympanometry, tone-audiograms Phoniatrician and clinical linguist for hearing and speech OM history: Groups defined by number of episodes of AOM during a 12 month period. Outcomes: • Phonology – based on material from *the phoneme test* and producing word pairs of the *auditory discrimination task*. • Grammar – from the *Ringsted material* • Interaction – *Thematic pictures* were used for this analysis • Grammatical and interactional analysis – based on sequential pictures and thematic pictures	Phonology % (#) had traces of deviance/deviant RAOM Healthy Age 4 69% (9/13) 76% (22/29) Age 7 83% (10/12) 54% (15/28) Grammar % (#) abnormal RAOM Healthy Age 4 0% (0/13) 0% (0/29) Age 7 0% (0/12) 0% (0/28) Interaction % (#) had traces of deviance RAOM Healthy Age 4 31% (4/13) 21% (6/29) Age 7 17% (2/12) 39% (11/28) RAOM=recurrent acute otitis media

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1623 Kaplan 1973	Prospective Cohort Study Study Quality Score (0–8): 6 (11001111) OM Diagnosis: A research nurse visited the cohort children 4 times a year during first 2 years and at least 2 times a year for the next 2 years. During each visit, information concerning middle ear abnormality was obtained and medical records of the time between visits were reviewed. Only OM described with otorrhea was used for analysis. OM Groups: Group 1: age of onset of first otorrhea at 0–1 year of age Group 2: age of onset of first otorrhea at age 2–10 years Group 3: no history of otorrhea N=489 N1=291 N2=83 N3=115	Time: Birth cohort (10/1960–12/1962); follow-up conducted between 9/1969–7/1971 Place: 28 Eskimo villages located in the Yukon and Kuskokwim River Delta areas of Southwestern Alaska Characteristics: Not described. Inclusion: • Age: Birth Cohort born between 10/1960–12/1962 and followed up between 9/1969–7/1971 • born in 25 villages w/in Alaska specified under place and/or residing in 3 additional villages at follow-up Exclusion: • Children with pure sensory hearing loss were excluded from analysis	Age: OM history: 0–1 year of life Outcome: at 10 years of age Examiner(s): Experienced audiologist for air and bone conduction measurements Two psychologists administered speech and intelligence testing OM history: Groups based on onset of first episode of otorrhea during 0–1 year; during 2–10 years or no history. Outcomes: • Wechsler Intelligence Scale for Children (WISC) – modified • Bender-Gestalt test • Draw-A-Person test Stratification Factor: Concurrent hearing status: No hearing loss or conductive component or hearing loss at 26+dB	**No Hearing Loss Group** 1st om at / Group 1 0–1y / Group 2 2–12y / Group 3 noOM N: 88 / 32 / 53 Verbal Score Mean: 80 / 79 / 81 Range: (50–100) / (55–105) / (55–105) Performance Score Mean: 100 / 100 / 102 Range: (70–140) / (70–125) / (70–135) **Conductive Component Group** 1st om at / Group 1 0–1y / Group 2 2–12y / Group 3 noOM N: 42 / 8 / 17 Verbal Score Mean: 70 / 80 / 76 Range: (45–90) / (60–100) / (55–90) Performance Score Mean: 92 / 100 / 95 Range: (55–125) / (85–110) / (75–115) **Hearing Loss 26+dB Group** 1st om at / Group 1 0–1y / Group 2 2–12y / Group 3 noOM N: 32 / 5 / 6 Verbal Score Mean: 72 / 71 / 80 Range: (55–100) / (55–85) / (60–100) Performance Score Mean: 95 / 94 / 101 Range: (55–125) / (70–120) / (70–115)

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
4651 Klein 1988	Prospective Cohort Study Study Quality Score (0–8): 4 (10100101) OM Diagnosis: By pediatricians using pneumatic otoscopy until age 3 and both pneumatic otoscopy and tympanometry in years 4 through 7. OM Groups: Group 1: Time spent with effusion <32 days during first 2 years of life Group 2: Time spent with effusion 33–108 days during first 2 years of life Group 3: Time spend with effusion >108 days during first 2 years of life. N=196 subjects Number of subjects by groups not given	Time: Enrollment of infants began in 1975 & continued for 2 yrs (7-year follow-up) Place: Private practices in Holliston and Framingham, MA, and an urban health center in East Boston Inclusion: • Random sample of children still active in the practices at age 7. Exclusion: • parental refusal • nonwhite race • English not the primary language spoken at home • Seizures • mental retardation, cerebral palsy • hearing loss >20db at time of testing (deferred until able to pass hearing test)	Age: OM history: first 3 years of life (recruited at <3 months of life) Outcome: at 7 years of age Examiner(s): Pediatricians for OM history. Not mentioned for speech and language testing. OM History: • Criteria for effusion – otorrhea, gas-liquid levels visible on otoscopy or marked reduction of mobility. Tympanometric criterion: type B curve. • Time spent with effusion – assigned a 'window' of 23 days to each observation of effusion, whether accompanied by signs of illness or not. It could be shortened or extended by multiple examinations. Outcome measure: • Wechsler Intelligence Scale for Children-Revised (WISC-R) • Goldman-Fristoe-Woodcock test • WUG test • Peabody Picture Vocabulary Test • Boston Naming Test • A recorded language sample • Metropolitan Achievement Test (reading and mathematics)	No raw data were given. Multivariate results were reported as follows: Least square (LS) mean and standard error (SE) of full scale WISC-R controlling for socioeconomic status and sex: Effusion Group Group 1 <32 d | Group 2 33–108 d | Group 3 >108 d LS mean 113.4 | 107.5 | 105.6 SE 1.41 | 1.49 | 1.34 Group 1 vs Group 2, p=0.005 Group 1 vs Group 3, p<0.0001 Group 2 vs Group 3, p=NS. The investigators report that 'similar results were found in both the performance and verbal scores. Speech and language measures significantly associated with otitis media during the first 3 years of life included aspects of speech production, speech perception, morphologic production, morphologic comprehension, lexical production, and syntactic production.

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1941 Mody 1999 Same study as 4780 Mody 1996 Same cohort as 1373, 2295, and 4728. See 1373, 2295, and 4728 for further comments.	Prospective Cohort Study Study Quality Score (0–8): 4 (11000101) OM Diagnosis: By trained and validated pediatric nurse practitioner (PNP) using a pneumatic otoscope under the supervision of a pediatric otolaryngologist. The PNP recorded a description of TM characteristics for each ear, using a 9-item otoscopic checklist and made the determination of "clear," "suspicious," or "positive" for OM. OM Group: Group 1: OM+: (Children who had 30% or more of the 13 first-year visits with otitis media bilaterally) Group 2: OM–: (Children who had 80% or more of the 13 first-year visits with normal middle ear findings bilaterally) N=14 subjects N1=7 N2=7	Time: Follow-up from birth to 9 years (actual dates not specified) Place: Subjects recruited from Longitudinal Infant Follow-up and Evaluation (LIFE) program of the Rose F. Kennedy Center, Albert Einstein College of Medicine, New York, NY Characteristics: • 71% high risk births • 64% male • 57% African-American • 36% Hispanic • 64% low SES • 36% middle SES Inclusion: • Age: 9 years (followed from birth) • English is the primary language spoken in the home • received all medical, audiological, and developmental assessments through the LIFE program at monthly evaluations from birth through first year • normal hearing and normal middle ear function on test day Exclusion: • neurologically compromised as an infant	Age: OM history: first year of life Outcome: at 9 years of age Examiner(s): Trained and validated pediatric nurse practitioner for OM history; Type of examiner(s) not specified for speech and language testing. OM history: Defined by pneumatic otoscopy findings during first year of life. Outcome measures: • Four nonsense syllables (/sβ /, /zβ /, /kβ /, /gβ /) presented auditorily • Test on identification and temporal order recall using 4 synthetic speech syllables, /ba/, /da/, /sa/ and /a/.	**Mean number of items correctly reported** OM+ OM– Serial Position 1 64 69 2 47 53 3 41 57 4 38 53 **Errors by type and number of phonetic features** OM+ OM– Transposition Errors 92% 95% Total errors 146 104 0 shared features 30 18 1 shared feature 35 21 2 shared features 81 65 **Mean (SD) number of errors on identification and temporal order judgment (TOJ)[a]** OM+ OM– 1. /ba/-/da/ Identification 1.0(1.0) 0.4(0.8) TOJ 400 0.7(1.5) 0.0(0.0) TOJ (100/50/10) 4.3(3.7) 1.1(0.9) 2. /sa/-/a/ Identification 2.1(2.3) 1.7(1.4) TOJ 400 3.1(2.5) 0.6(0.8) TOJ (100/50/10) 6.7(5.7) 3.6(3.2) [a] From article 4780.

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
4675 Owen 1996	Prospective Cohort Study Study Quality Score (0–8): 4 (11000101) OM Diagnosis: By trained technicians using automated screening tympanometers with a 226 Hz probe tone. Acoustic reflectivity was also measured using acoustic otoscope at 30% of visits. OME diagnoses were based on type B tympanogram or >= 5 acoustic reflectivity or visible purulent otorrhea without an otoscope. OM Groups: Group 1: High OME (children who had extended OME across first three years of life) Group 2: Early OME (children who had OME which peaked from 0–6 months of age) Group 3: Later OME (children who had OME which peaked from 6–12 months of age) Group 4: Low OME (children who experienced relatively low levels across first 3 yrs of life) N=294 subjects Sample size for each group not specified.	Time: Infants enrolled in program between 1984–1989 (5 yr follow-up) Place: Texas Characteristics: • 49% male • 56% Euro-American • 30% African-American • 14% English-speaking Hispanic • wide socioeconomic range • Half breast fed at birth • 41% had cigarette smoke exposure • 19% mother smoked • 63% attended day care Inclusion: • Age: 5 years (follow-up from birth) • healthy-term infant from English-speaking family Exclusion: None	Age: OM history: 0–3 years of life Outcome: at 5 years of age Examiner(s): Trained technicians for OM history Trained graduate students for language and cognitive testing. OM history: If 2 consecutive visits showed OME, the intervening days were counted as days with OME. If one visit showed OM and the next normal status, or vice versa, half of the intervening days were counted as days with OME. OME duration was defined as the proportion of time a subject spent with OME (total OME days divided by total days) in the period examined. Outcome measures: • Test for Auditory Comprehension of Language-Revised (TACL-R) • Auditory discrimination subtests of the Carrow Auditory Visual Abilities Test (CAVAT) • Goldman-Fristoe test for articulation (G-F) • Carrow Elicited Language Inventory (CELI) • Stanford-Binet Fourth Edition for cognition Covariates: Gender, Ethnicity, Birth rank, Day care attendance, Duration of breast feeding, Cigarette smoke exposure, SES, Educational stimulation at home, Mother's intelligence	Raw results were not reported. The following findings were reported: 1. CELI, a measure of child's productive control of grammar, showed a significant inverse relation to duration of OME [t(266)=–2.62; p=0.01]. 2. Goldman-Fristoe, a measure of articulation errors, OME amount was positively related to articulation errors [t(213)–6.42;p=0.01] 3. CAVAT yielded significant relations for auditory blending and auditory discrimination in quiet modified by ethnicity and breast feeding. 4. TACL-R total and word class scores were also related to OME cluster group. 5. High OME cluster group tended to be most adversely affected, but all relations were moderated, chiefly by ethnicity, breast feeding and home stimulation scores. 6. Association of duration of OME in first 18 months of life and Stanford Binet composite scores: r=–0.11, p<0.05. 7. Association of duration of OME and nonverbal reasoning and visualization factor scores: r=0.13, p<0.05) 8. Associations in 5 and 6 were not substantiated when effects of environmental and demographic variables were controlled in multivariate analysis.

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
2135 Paul 1993	Retrospective-Prospective Cohort Study Study Quality Score (0–8): 3 (00100101) OM Diagnosis: Based on parental reporting. OM Group: Middle ear involvement was defined as either the placement of myringotomy tubes or the presence of six or more ear infections treated by a physician before the second birthday, by parent report. Group 1: OM+: defined by middle ear involvement, stratified by normal or late talkers Group 2: OM–: lack of such middle ear involvement, also stratified by normal or late talkers. N=44 subjects N1=20 (8 normal and 12 late talkers) N2=24 (13 normal and 11 late talkers)	Time: Not specified Place: Portland Oregon. Subjects are subset of those participating in the Portland Language Development Project (PLDP), a longitudinal study of outcomes of early expressive language delay. Characteristics: • all passed speech reception screenings • 15–33% of each subgroup had abnormal tympanograms Inclusion: • Age: 20–34 at entrance • subjects recruited from children participating in the Portland Language Development Project (PLDP) Exclusion: None	Age: OM history: first 2 years of life Outcome: at 3 and 4 years of age Examiner(s): Certified audiologist in sound-treated booth. OM History: Based on parental reports. Outcome measures: • Goldman-Fristoe Test of Articulation at age 3 • Word Articulation subtest of the Test of Language Development-Primary at age 4. • Sample of free speech collected during a 15-minute play interaction between mother and child at ages 3 and 4. • Speech samples were used to compute mean length of utterance (MLU) in morphemes for each sample. Stratification groups: Children were placed in normal versus late talking (LT) groups on the basis of their expressive vocabulary size as reported by parents on the Language Development Survey (LDS) at intake into the PLDP.	**MLU – Mean (SD)** OM+ OM– Normal Talkers n=8 n=13 Age 3 4.26 (1.56) 4.05 (0.85) Age 4 4.82 (0.32) 4.55 (0.89) Late Talkers n=12 n=11 Age 3 2.42 (0.60) 2.61 (0.79) Age 4 4.08 (1.05) 3.91 (1.09) **Goldman-Fristoe %tile – Mean (SD)** OM+ OM– Normal Talkers[a] n=8 n=13 Age 3 58.9 (33.0) 62.3 (32.4) Age 4 57.1 (26.0) 54.5 (37.8) Late Talkers[a] n=12 n=11 Age 3 12.3 (10.8) 18.9 (19.8) Age 4 17.8 (19.0) 31.6 (28.6) [a] Children who at 20 to 34 months of age used fewer than 50 words were assigned to the LT group. About one-third of the LT subjects showed some delay in receptive language abilities, but they generally scored within 6 months of age level on the Reynell Receptive Language Scales.

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
4657 Roberts 1986 4656 Roberts 1989 Same study subjects as in 4806/ 3118, 3117, and 4319. See 4806/ 3118, 3117, and 4319 for further comments.	Prospective Cohort Study Study Quality Score (0–8): 6 (10101111) OM Diagnosis: By pediatricians and pediatric nurse practitioners based on pneumatic otoscopy. 60% time tympanometry was used to corroborate the OME diagnosis. OM Groups: Based on total OME duration in days during first 3 years of life. Group 1: Days with total OME between 0–87 days. Group 2: Days with total OME between 88–181 days. Group 3: Days with total OME over 181 days. N=61 subjects up to 6 years N1=20 N2=20 N3=21 N=44 subjects up to 8 years N1=15 N2=14 N3=15	Time: 9/1972–12/1984 (follow-up from birth to third year of school) Place: Frank Porter Graham Child Development Center, U of North Carolina at Chapel Hill, NC Characteristics of 61 subjects • African American • 57% boys • Average of 10.5 years of education of mother • Mother's mean IQ score was 84. • Hearing was within normal limits at time of testing Inclusion: • Age: 3.5–6 years of age at followup • participant in longitudinal on-site research day-care program • attended the Frank Porter Graham Child Development Center, 5 full days/wk, 50 wks/yr through 5 yrs of age • identified at birth as biologically normal • Classified as at risk for poor school performance Exclusion: • incomplete data	Age: OM history: 2 months–3 years; Outcome: at 3.5–8 years of age Examiner(s): Pediatricians and pediatric nurse practitioners for clinical history Psychologists for measurements of child development. OM History: Duration of each episode of unilateral and bilateral was calculated by subtracting the date of onset of OME from the resolution date. Days of total Outcome measures: Cognitive capacity: • Stanford-Binet Intelligence Scale • McCarthy Scale of Children's Abilities (McCarthy) • Wechsler Preschool and Primary Scale of Intelligence (WPPSI) • Wechsler Intelligence Scale for Children-Revised (WISC-R) verbal scale • Classroom Behavior Inventory (CBI) verbal intelligence Academic achievement: • Peabody Individualized Achievement Test (PIAT) • Woodcock-Johnson Psychoeducational Battery (WJPB)	**Age 3.5–4 years** Mean (SD) McCarthy: Group 1 (n=20) / Group 2 (n=20) / Group 3 (n=21) Cognitive: 103 (12) / 104 (10) / 100 (8) Verbal: 52 (7) / 54 (7) / 52 (5) Stanford-Binet Intelligence: (n=20) / (n=20) / (n=19) 101 (11) / 98 (8) / 99 (11) **Age 4.5 years** Mean (SD) McCarthy: Group 1 (n=19) / Group 2 (n=20) / Group 3 (n=19) Cognitive: 100 (13) / 103 (9) / 97 (10) Verbal: 53 (9) / 55 (6) / 52 (8) **Age 5 years** WPPSI Intelligence: (n=19) / (n=20) / (n=20) Full-scale: 102 (12) / 101 (10) / 98 (11) Verbal: 104 (11) / 101 (9) / 97 (12) Performance: 100 (13) / 101 (11) / 100 (11) **Age 8 years** Groups by OME days <82d (n=15) / 83–157d (n=14) / >157d (n=15) Intelligence – Verbal IQ WISC-R: 98 (13) / 96 (14) / 94 (10) CBI: 14 (5) / 12 (3) / 12 (3)

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
4806 Roberts 1988 3118 Roberts 1988 Same study subjects as in 4657/ 4656, 3117, and 4319. See 4657/ 4656, 3117, and 4319 for further comments	Prospective Cohort Study Study Quality Score (0–8): 4 (10100101) OM Diagnosis: By pediatricians and pediatric nurse practitioners based on pneumatic otoscopy. 60% time tympanometry was used to corroborate the OME diagnosis. OM Groups Based on total OME duration in days during first 3 years of life. Duration of each episode of unilateral and bilateral was calculated by subtracting the date of onset of OME from the resolution date. Group 1: lower third–OME days (8–143 days) Group 2: middle third–OME days (144–269 days) Group 3: upper third–OME days (314–931 days) N=55 children tested for speech at different ages. 12 tested once; 7 tested 2 times; 19 tested 3 times; 6 tested 4 times and 11 tested 5 times. N=34 in the language study	Time: 1975–1986 study period (8 yr follow-up reported); speech and language measured between 1982–1986 Place: Frank Porter Graham Child Development Center, Un. of North Carolina at Chapel Hill, NC Characteristics: • 93% black, 7% white • 67% boys • All passed hearing screen and had type A or C tympanogram. Inclusion: • Age: 8 years (followed from <3 months) • Attended Frank Porter Graham Child Development Center for 5 days/wk, 50 wks/yr until entry into kindergarten • Socioeconomically disadvantaged • identified at birth as biologically normal • Classified as at risk for poor school performance • Tested for speech and language between 1982–1986 Exclusion: None	Age: OM history: 2 months–3 years; Outcome: at 2.5–8 years of age Examiner(s): Pediatricians and pediatric nurse practitioners for clinical history. Psychologists for measurements of child development. Master's level speech-language pathologists and graduate research assistants collected speech data. OM History: Days of total OME was analyzed both as a continuous and categorical variable. Outcome measures: Speech • Goldman-Fristoe Test of Articulation to which 6 additional words were added. Transcription and Phonology • CER – total number of consonants in error • TPP – total number of phonological processes Language – 15-minute elicited conversational speech sample of language play situation. • Number of words per communication unit (cu) • Dependent clauses per cu • Ratio of total different words to total cu • Ratio of total different conjunction words (excluding 'and' to total cu • Number of utterances per turn.	Spearman coefficient of correlation (R) between number of phonologic processes and total OME days Age(yr) N R P-Value 3 27 –0.04 0.85 4 25 0.29 0.16 5 24 0.31 0.15 6 30 0.29 0.13 7 29 0.18 0.34 8 19 0.10 0.69 Spearman coefficient of correlation (R) between language measures and total OME days at age 5 (n=34) Measure R P-Value Words/cu –0.04 0.83 Dependent Clauses/cu 0.17 0.33 Different Words/cu –0.03 0.89 Different Conjunctions/cu 0.20 0.26 Different Pronouns/cu 0.09 0.62 Regression analysis revealed that (a) children with more OME tended to have a greater total number of phonological processes and (b) Unilateral OME(0–3) provided independent prediction of median TPP given Bilateral OME(0–3), but that Bilateral OME(0–3) was a nonsignificant predictor given Unilateral OME(0–3).

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
3117 Roberts 1991 Same study subjects as in 4657/ 4656, 4806/3118and 4319. See 4657/ 4656, 4806/ 3118, and 4319 for further comments.	Prospective Cohort Study Study Quality Score (0–8): 2 (00000101) OM Diagnosis: By pediatricians and pediatric nurse practitioners based on pneumatic otoscopy. 60% time tympanometry was used to corroborate the OME diagnosis. OM Groups: Based on total OME duration in days during first 3 years of life. Duration of each episode of unilateral and bilateral was calculated by subtracting the date of onset of OME from the resolution date. N=63 subjects Subjects were stratified into two groups SES status: Group 1: LSES (Children from families of lower socioeconomic status) Group 2: MSES (Children from families of middle socioeconomic status) N1=33 N2=30	Time: Cohort of 8–12 children entered every 1–2 years between 1978 and 1985 Place: Attended the Frank Proter Graham Child Development Center (FPG), a research day-care center, University of North Carolina at Chapel Hill Characteristics: LSES: male 64%, black 88%, single mom 82% MSES: male 50%, black 33%, single mom 3% • all passed hearing screen, had type A or C tympanogram, and did not have OME Inclusion: • Age: 6 years (follow-up from birth) • attended the Frank Porter Graham Child Development Center • identified at birth as biologically normal • Classified as at risk for poor school performance Exclusion: None	Age: OM history: 2 months–3 years; Outcome: at 4.5–6 years of age Examiner(s): Pediatricians and pediatric nurse practitioners for clinical history. Psychologists for measurements of child development. Master's level speech-language pathologists and research assistants for speech data. OM History: Days of total OME was analyzed both as a continuous and categorical variable. Outcome measures: Language tests Age 4–5 years • Miller-Yoder Language Comprehension Test (M-Y) • Bankson Language Screening Test (Bankson) • Clinical Evaluation of Language Functions (CELF) • Peabody Picture Vocabulary Test-Revised (PPVT-R) At Age 5, 15-minute elicited sample of language play situation • Words/communication unit (cu) • Dependent clauses per cu • Ratio: total different words to total cu • Ratio: total different conjunction words (excluding 'and') to total cu • Ratio: different pronouns to total cu • Number of utterances per turn.	Spearman's correlation coefficient (R) between language outcomes and total OME days during first 3 years of life LSES: Bankson N=25 R=0.13; M-Y N=25 R=0.08; PPVT-R N=24 R=–0.08; CELF receptive N=25 R=0.25; CELF expressive N=24 R=0.23 MSES: Bankson N=22 R=0.02; M-Y N=22 R=0.14; PPVT-R N=20 R=0.03; CELF receptive N=21 R=0.24; CELF expressive N=21 R=0.17 *Language Sample* Words/cu: LSES N=28 R=0.05; MSES N=26 R=–0.22 Dependent Clauses/cu: LSES N=28 R=0.04; MSES N=26 R=–0.02 Different Words/cu: LSES N=28 R=0.26; MSES N=26 R=0.00 Different Conjunctions/cu: LSES N=28 R=0.02; MSES N=26 R=0.13 Different Pronouns/cu: LSES N=28 R=0.24; MSES N=26 R=–0.12

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings

4319
Roberts
1995
 	
This is a continuation of study by Roberts.
4657
4656
 	
Same study subjects as in 4657/ 4656, 4806/ 3118 and 3117
 	
See 4657/ 4656, 4806/ 3118, and 3117 for further comments.
	Prospective Cohort Study
	
	Study Quality Score (0–8):
	3 (00000111)
	
	OM Diagnosis:
	By pediatricians and pediatric nurse practitioners based on pneumatic otoscopy. 60% time tympanometry was used to corroborate the OME diagnosis.
	
	OM Groups: Based on total OME duration in days during first 3 years of life.
	
	Group: African American children enrolled as infants in research child-care program (follow-up to 12 years of age)
	
	N=56 of 61 subjects reported in 4657 and 4656.
		Time: Cohort assembled 9/1972–12/1984, follow-up from birth to 12 years of age)
		
		Place: Frank Porter Graham Child Development Center, U of North Carolina at Chapel Hill, NC
		
		Characteristics of 61 subjects:
		• African American
		• 57% boys
		• Average of 10.5 years of education of mother
		• Mother's mean IQ score was 84.
		• Hearing was within normal limits at time of testing
		
		Inclusion:
		• Age: 12 years of age at followup
		• participant in longitudinal on-site research day-care program
		• attended the Frank Porter Graham Child Development Center, 5 full days/wk, 50 wks/yr through 5 yrs of age
		• identified at birth as biologically normal
		• Classified as at risk for poor school performance
		
		Exclusion:
		• incomplete data
			Age:
			OM history: 2 months–3 years; Outcome: at 12 years of age
			
			Examiner(s):
			Pediatricians and pediatric nurse practitioners for clinical history
			
			Psychologists for measurements of child development.
			
			OM History:
			Duration of each episode of unilateral and bilateral was calculated by subtracting the date of onset of OME from the resolution date.
			
			Summary measures:
			• No. OME – total number of OME episodes
			• ML/OME – mean length in number of days of the OME episodes
			• Duration OME – number of days with either bilateral or unilateral OME.
			
			Outcome measures:
			Cognitive capacity:
			• Wechsler Intelligence Scale for Children-Revised (WISC-R)
			  • Full-scale IQ
			  • verbal IQ
			  • performance IQ
			  • Freedom from Distractibility Factor
			
			Academic performance:
			• Woodcock-Johnson Psychoeducational Battery
			
			Behavior
			• Child Behavior Checklist
				Correlation coefficient, R, between language outcomes and total OME days during first 3 years of life
				
				At Age 12
				Measure N R1 R2
				Intelligence-WISC-R
				Full scale IQ 56 –0.17 –0.12
				Verbal IQ 56 –0.10 –0.08
				Performance IQ 56 –0.18 –0.13
				
				Where R1 is the simple correlation coefficient and R2 is the standardized regression coefficient adjusted for gender and home environment.
				
				The multivariate tests of the relationships between the ranks of No. OME(0–3) and ML/OME(0–3) during early childhood and the Full-Scale, Verbal, and Performance IQ were not significant either when the HOME or gender were considered or ignored.

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
2295 Ruben 1997 Same study cohort as 1373, 1941, and 4728 by Gravel See 1373, 1941, and 4728 for further comments.	Prospective Cohort Study Study Quality Score (0–8): 5 (11001110) OM Diagnosis: Pediatric nurse practitioners completed pneumo-otoscopic examinations. OM Groups: Group 1: Otitis positive children – when bilateral OM was detected at 30% or more of the baby's first year visits. Group 2: Otitis negative children – when middle ear status was rated as normal in both ears during 80% or more of the first year visits. N=30 subjects N1=18 N2=12	Time: not specified Place: Clinical Research Center for Communicative Disorders (which draws subjects from the LIFE Program of the R.F. Kennedy Center, Albert Einstein College of Med. Bronx, NY Characteristics: • low socioeconomic urban neighborhoods based on Hollingshead's (1975) index Inclusion: • Age: 9 years • born in the same urban hospital • auditory testing conducted during first year of life and at 9 years of age Exclusion: None	Age: OM history: 0–1 of life Outcome: 2–9 years of age Examiner(s): Not specified OM history: OM groups were defined by otoscopic histories. Outcome measures: 18 measures of communicative function was measured. Data were summarized by calculating the mean score for OM– group and then comparing the score of each child in the OM+ group with that of the mean for the OM– group. Each OM+ score that was >= the mean of the OM– group was counted and the percent of children in the OM+ group who had higher scores than the OM– was calculated for each measure. Differences between the groups were considered meaningful when less than 40% of the OM+ children in the group exceeded the score of the OM– group.	Percent OM+ children performed more poorly than the OM- in the following measure: Measure % Expressive language at 1 yr 0% Expressive language at 2 yr 8% Expressive language at 4 yr 25% Receptive language competition at 4 yr 38% Speech in competition ratio as measured adaptively at 4 yr 30% Comprehension of grammatical structures at 6 yr 11% School readiness at 6 yr 0% Recall of a narrative at 9 yr 14%

Evidence Table 2. Early Life OM and Long-Term Speech and Language Development (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
2583 Teele 1990 Same cohort as 2579	Prospective Cohort Study Study Quality Score (0–8): 6 (10101111) OM Diagnosis: By physicians and nurse practitioners using a standard sealed pneumatic otoscope. To resolve ambiguous diagnoses, otoadmittance was used sporadically at 0–3 years of age, but frequently at 4–7 years. Group: Children followed prospectively from birth until age 7 years and whose cognitive and linguistic abilities were assessed at age 7 N=194 subjects	Time: Enrollment of infants began 6/1975 (follow-up during first 7 yrs of life) Place: Subjects drawn from larger cohort of children enrolled at an urban health center in East Boston and a private practice in the Holliston/Framingham area of MA Characteristics: • 54% male • 57% high SES • 69% from private practice Inclusion: • Age: 7 years (followed from birth) • Caucasian Exclusion: • seizures • mental retardation • cerebral palsy • two languages spoken in home • non-white	Age: OM history: 0–2 years Outcome: at 7 years of age Examiner(s): Physicians and nurse practitioners for clinical history. Experienced psychometricians for intelligence and achievement tests Certified speech clinicians administered speech and language tests. OM history: Used the number of days with MEE by age 3 years as a predictor variable. Unless documented to be shorter, each episode of MEE lasted 29 days. Outcome measures: Cognitive assessment • WISC-R Speech and language • Subsample of Goldman-Fristoe and Goldman-Fristoe-Woodcock Tests • WUG test • Peabody Picture Vocabulary Test • Boston Naming Test • Complex sentences and set of pictures for story telling Academic achievement • Metropolitan Achievement Test	Cognitive ability (mean IQ by WISC-R) at 7 years of age after adjusting for SES status and gender. Time with MEE <30d (n=58) 30–129d (n=77) >130d (n=59) Full Scale 113 108 105 Verbal 112 106 106 Performance 112 108 104 Linguistic abilities (mean score) after adjusting for SES status and gender Time with MEE <30d (n=58) 30–129d (n=77) >130d (n=59) Goldman-Fristoe-Woodcock Articulation 106 66 56 Discrimination 34 30 28 WUG Receptive 1.5 1.8 2.2 Expressive 2.4 2.0 1.9 Morphologic markers 31 26 20 Lexical use 1.5 1.5 1.5 Speech sample score 83 80 78

Evidence Table 3: Early Life OM and Long-Term Hearing

[The eight components of study quality score are: study cohort clearly defined; subjects assembled at a uniform time point; pathway of subject entry clearly described; complete follow-up achieved; withdrawals/drop-outs described; objective outcomes used; outcome assessment blinded; and extraneous factors adjusted. 1 indicates presence and 0 indicates absence.]

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
147 Sorri 1995	Prospective Cohort Study Study Quality Score (0–8): 4 (01001101) OM diagnosis: By whom not specified; How not specified. OM Groups: Group 1: No otitis media until 7 years of age (NoOM) Group 2: Recurrent otitis media (>=4 episodes of OM until the age of two) (RAOM) Group 3: Secretory otitis media until the age of two (SOM) N=99 N1=35 N2=51 N3=13	Time: Birth cohort 1985–1986 Place: Health care centers, hospital and private surgeries in Northern Finland Inclusion: • Age: Born in Northern Finland between 1985–1986 (examined through age 2 and at 7 years) • history of OM meticulously controlled up until the age of 2 yrs Exclusion: • middle ear effusion at time of audiometry at age 7 (14 excluded) • 6 children had meningitis, 1 totally deaf child, 1 w/atresia of the ear canal, 5 lacked cooperation, 5 lacked information after age 2 years	Age: OM history: 0–2 years of life Outcomes: At age 7 years Examiner(s): Not specified OM history: OM history groups defined by number of recurrent otitis media or serous otitis media. Outcomes: • Pure tone average, type not specified • Mean air-conduction (AC) thresholds	**Number and percent children exceeding 20 dB (hearing loss)** NoOM RAOM SOM 7/35 28/51 5/13 20% 55% 38% **Pure tone (dB), right ear: mean±SD** NoOM RAOM SOM 3.82±3.50 4.88±4.42 6.44±6.27 **Pure tone (dB), left ear: mean±SD** NoOM RAOM SOM 3.24±2.21 3.63±2.65 6.54±6.56

Evidence Table 3. Early Life OM and Long-Term Hearing (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1255 Fischler 1985	Retrospective-prospective study Study Quality Score (0–8): 5 (11001110) OM Diagnosis: By medical record review of documented physician's clinical diagnosis. OM Groups: Group 1: HR/REC (HR-High risk status,REC-recurrences of OM attacks after age 2) ~ >=2 attacks by age 2 yr and >=3 attacks after age 2 yr Group 2: HR/NREC ~ >=2 attacks by age 2 yr and <3 attacks after age 2 yr Group 3: NHR/NREC ~ <2 attacks by age 2 yr and <3 attacks after age 2 yr N=167 N1=33 N2=63 N3=71	Time: Start date~7/1974 Place: Four Indian reservations in Arizona Characteristics: • 50% had family income under $5000 per year • 37% had housing below average • 13% primarily Apache speaking at home Inclusion: • Age: 6–8 years • healthy Apache Indian children who had been followed since birth Exclusion: • moved during study period • absent or not tested because of time constraints • medical reasons (specific conditions not specified)	Age: OM history: 0–2 years of life Outcomes: At 6–8 years Examiner(s): • Unspecified physicians for diagnosis of otitis media; • School nurse for hearing tests; • Pediatrician/otolaryngologist for otoscopic examinations; • Certified speech pathologists for language testing. OM history: Groups defined by number of OM episodes by age 2 Outcomes: All children received a hearing screen using a recently calibrated pure tone audiometer (dB hearing level American National Standards Institute [ANSI] 1969). Children whose hearing was better than the following thresholds were considered normal: 25 dB at 500 Hz; 20 dB at 1,000 Hz; 20 dB at 2,000 Hz; 25 dB at 4,000 Hz; 25 dB at 6,000 Hz; Otherwise, abnormal or hearing loss.	**Number and percent of children with hearing loss** Group 1 Group 2 Group 3 OM risk High Moderate Low 5/33 4/63 1/71 15% 2% 1%

Evidence Table 3. Early Life OM and Long-Term Hearing (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1373 Gravel 1992	Prospective Cohort Study Study Quality Score (0–8): 5 (11001101) OM Diagnosis: Pediatric nurse practitioners completed pneumo-otoscopic examinations. OM Groups: Group 1: Otitis positive children – when bilateral OM was detected at 30% or more of the baby's first year visits. (Also, had poorer auditory sensitivity by click ABR.) Group 2: Otitis negative children – when middle ear status was rated as normal in both ears during 80% or more of the first year visits. N=23 N1=10 N2=13	Time: not specified Place: Clinical Research Center for Communicative Disorders (which draws subjects from the LIFE Program of the R.F. Kennedy Center, Albert Einstein College of Med. Bronx, NY Characteristics: • low socioeconomic urban neighborhoods based on Hollingshead's (1975) index • none had sensorineural hearing loss Inclusion: • Age: 4 years (measures taken during first year of life as well) • all children were enrolled at 40 wks postconceptional age for follow-up through the LIFE Program • either high-risk infant graduates of a NICU or FT graduates of the well-baby nursery of Jacobi Hospital • English reported as the primary language spoken Exclusion: • neurologically compromised	Age: OM history: 0–1 of life Outcome: at 4 years of age Examiner(s): Not specified OM history: OM groups were defined by otoscopic histories. Outcomes: Pediatric Speech Intelligibility (PSI) sentence (S) items/competing messages (CM) ratio Pure-tone averages – based on a conventional clinical staircase procedure (descending-ascending). Thresholds for each ear of all subjects were obtained at octave frequencies from 500 through 4000 Hz (at a minimum) (ANSI, 1969).	**PSI S/CM (dB)** OM+ OM– Mean±SD -6.8±2.8 -9.7±2.6 **Pure-tone (dB), right ear** OM+ OM– Mean±SD 15.8±4.1 14.9±3.2 **Pure-tone (dB), left ear** OM+ OM– Mean±SD .8±4.1 14.6±3.3

Evidence Table 3. Early Life OM and Long-Term Hearing (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1435 Harsten 1993	Prospective Cohort Study Study Quality Score (0–8): 6 (11101110) OM Diagnosis: AOM was diagnosed by otomicropscopy, performed by an otolaryngologist and defined as an acute episode of earache in a child, with red bulging eardrum(s) or purulent discharge, occasionally febrile and with signs of upper respiratory tract infection. OM Groups: Group 1: Children with recurrent AOM (RAOM defined as at least six episodes of AOM during a 12-month period) during the first 3 years of life Group 2: Children without any AOM episode during the first 3 years of life N= 42 N1=13 N2=29	Time: not specified Place: University Hospital of Lund, Sweden Characteristics: • 45% first-borns • 71% mother educated at college level Inclusion: • Age: Birth cohort • monolingual, Swedish children, born at University Hospital of Lund Exclusion: None	Age: OM history: 0–3 year of life Outcomes: at 4 and 7 year of age Examiner(s): Otolaryngologist for otomicroscopy, tympanometry, tone-audiograms Phoniatrician and clinical linguist for hearing and speech OM history: Groups defined by number of episodes of AOM during a 12-month period. Outcomes: Tone-audiometry: abnormal defined as >=25 dB threshold at any frequency. (Tone-audiograms were recorded at frequencies from 125 to 8000 Hz.)	**Tone-audiometry at age 4** RAOM Healthy Abnormal 3/26 4/58 hearing 12% 7% **Tone-audiometry at age 7** RAOM Healthy Abnormal 2/24 3/56 Hearing 8% 5%

Evidence Table 3. Early Life OM and Long-Term Hearing (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
1623 Kaplan 1973 Same study subjects as 2233 which reported 3–5 year follow-up. This study reported 10-year follow-up.	Prospective Cohort Study Study Quality Score (0–8): 5 (11001110) OM Diagnosis: A research nurse visited the cohort children 4 times a year during first 2 years and at least 2 times a year for the next 2 years. During each visit, information concerning middle ear abnormality was obtained and medical records of the time between visits were reviewed. Only OM described with otorrhea was used for analysis. OM Groups: Group 1: age of onset of first otorrhea at 0–1 year of age Group 2: age of onset of first otorrhea at age 2–10 years Group 3: no history of otorrhea N=489 N1=291 N2=83 N3=115	Time: Birth cohort (10/1960–12/1962); follow-up conducted between 9/1969–7/1971 Place: 28 Eskimo villages located in the Yukon and Kuskokwim River Delta areas of Southwestern Alaska Characteristics: Not described. Inclusion: • Age: Birth Cohort born between 10/1960–12/1962 and followed up between 9/1969–7/1971 • born in 25 villages w/in Alaska specified under place and/or residing in 3 additional villages at follow-up Exclusion: • Children with pure sensory hearing loss were excluded from analysis	Age: OM history: 0–1 year of life Outcome: at 10 years of age Examiner(s): Experienced audiologist for air and bone conduction measurements Two psychologists administered speech and intelligence testing OM history: Groups based on onset of first episode of otorrhea during 0–1 year; during 2–10 years or no history. Outcomes: Air and bone conduction measurements were made using a portable audiometer recently calibrated to the ISO 1964 standard. All children were tested at frequencies of 250, 500, 1000, 2000, 4000, and 8000 Hz to evaluate air conduction and at frequencies of 500, 1000, 2000, and 4000 Hz to evaluate bone conduction. • Normal hearing: 0 to –25 dB; • Normal hearing with a conductive component: 0 to –25 dB with a 15 dB air-bone gap; • Hearing loss: –26 dB or greater. (Children with pure sensory hearing loss were excluded from analysis.)	**Hearing Loss at 10 years of age** 1st om at: Group 1 0–1y; Group 2 2–12y; Group 3 noOM Conductive hearing loss (air-bone gap) 42/162 8/45 17/76 28% 18% 22% Hearing loss at 26+dB 32/162 5/45 6/76 20% 11% 8%

Evidence Table 3. Early Life OM and Long-Term Hearing (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
2233 Reed 1967 Same study population as 1623. This article reported findings at 3–5 years. Article 1623 reported findings at 10 years of age.	Prospective Cohort Study Study Quality Score (0–8): 4 (11000101) OM Diagnosis: A research nurse visited the cohort children 4 times a year during first 2 years and at least 2 times a year for the next 2 years. During each visit, information concerning middle ear abnormality was obtained and medical records of the time between visits were reviewed. Only OM described with otorrhea was used for analysis. OM Groups: Defined by frequency of otorrhea episodes during first 2 years of life. Group 1: None Group 2: <1 episode per year Group 3: 1 or more episode per year N=378 N1=198 N2=180	Time: Two birth cohorts (born between 10/1960–12/1962); follow-up conducted between Sept.-Dec.,1965 Place: Testing conducted in schoolhouses or in a community building in 24 of the original 27 Eskimo villages located in the Yukon and Kuskokwim River Delta areas of SW Alaska where the cohort was born Characteristics: Not described. Inclusion: • Age: Follow-up from birth to 4 years (0–47 months) • Born in one of 27 Eskimo villages between 10/1960–12/1962 and available for testing at follow-up • audiometric tests completed Exclusion: • Children with pure sensory hearing loss was excluded from analysis	Age: OM history: first 2 years of life Outcome: at 3–5 years of age Examiners: Audiologist OM history: Groups defined by frequency of otorrhea episodes during first 2 years of life: none, <1 per year of risk, and 1+ per year of risk Outcomes: Air and bone conduction measurements were made using the Zenith ZA-100-T Diagnostic Portable Audiometer, calibrated to the ISO 1964 standards. Children were classified according to averages of pure tone air hearing thresholds at 500, 1000, and 2000 CPS: • normal: 0 to 25 dB • mild impairment: 26 to 40 dB • moderate to severe impairment: 41 to 70+ dB	**Mild and Severe Hearing Loss at 3–5 years of age** OM history, first 2 years of life none: 21/143, 15% <1/yr: 48/139, 35% 1+/yr: 47/96, 49%

Evidence Table 3. Early Life OM and Long-Term Hearing (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
2309 Ryding 1997	Prospective Cohort Study Study Quality Score (0–8): 5 (01100111) OM Diagnosis: AOM was diagnosed by otomicroscopy, performed by an otolaryngologist and defined as an acute episode of earache in a child, with red bulging eardrum(s) or purulent discharge, occasionally febrile and with signs of upper respiratory tract infection. OM Groups: Group 1: RAOM Children (>=6 episodes of purulent AOM during a 12-month period) Group 2: Healthy Children (no AOM and <6 other RTI episodes during the study period) N=33 N1=12 N2=21	Time: Birth cohort (born between 11/1982–2/1984); 10 yr follow-up Place: University Hospital, Lund, Sweden Characteristics: • 45% first-borns • 71% mother educated at college level Inclusion: • 10 years old • member of this birth cohort Exclusion: • moved out of Lund	Age: OM history: 0–3 years of life Outcome: at 10 years of age Examiner(s): Audiologist Instrument: tone audiometry GSI 16 audiometers OM history: Groups defined by number of recurrent AOM. Recurrent AOM was defined as >= 6 episodes of purulent AOM during a 12-month period. Outcomes: Children were tested using GSI 16 audiometers with TDH39 earphones calibrated according to ISO 389. Air conduction hearing levels for both ears were individually determined at the frequencies 1000, 1500, 2000, 3000, 4000, 6000, 8000, 500, 250 and 125 Hz, tested in that order. The test was conducted according to the ascending method. Pure tone average (PTA) values for frequencies 500, 1000 and 2000 Hz (PTA: 0.5-2) and 500, 1000, 2000 and 4000 Hz (PTA: 0.5-4) were calculated for each ear.	**Median level in dB, right ear** RAOM Healthy PTA(0.5-2) 5 5 PTA(0.5-4) –2 –2 **Median level in dB, left ear** RAOM Healthy PTA(0.5-2) 7 3 PTA(0.5-4) –2 0

Evidence Table 3. Early Life OM and Long-Term Hearing (Continued)

Record# Author Year	Study Quality OME Definition Group(s) and Sample Size	Time/Place Subject Characteristics Inclusion/Exclusion Criteria	Risk Groups Predictors Outcome Measures	Findings
2854 Zargi 1992	Retrospective prospective cohort study Study Quality Score (0–6): 2 (10000100) OME definition: Not specified Group 1: Experimental – Children treated for recurrent acute unilateral or bilateral suppurative otitis media at 0–2 years of age. Group 2: Control – Children who experienced <=1 episodes of OM in the first 2 years of life) N=62 N1=33 N2=29	Time: not specified Place: University Department of Otorhinolaryngology and Head and Neck Surgery, Ljubljana, Slovenia Inclusion: • Age: 8–10 years • attended similar pre-school, kindergarten, and elementary school programs • no evidence of mental retardation nor behavioral or emotional disorders Exclusion: None	Age: Followup examination at 8–10 years of age Examiner(s): Otolaryngologist OM history: During first two years of life. Outcomes: At age 8–10 years pure-tone audiometry was performed. The criterion for normal hearing was based on the air-conduction threshold results, which had to be at or above the level of 10 dB. Isolated audiometric dips (notches) to 15 dB at not more than one high frequency (6kHz or 8kHz) were not considered to be sensorineural hearing loss.	Hearing loss Group 1 Group 2 OM+ OM– Sensorineural 10/33 4/29 30% 14% Conductive or mixed 8/33 0/29 24% 0% Both types 18/33 4/29 55% 14%

Evidence Table 4: Diagnostic Methods

[The six components of Study Quality Score are: appropriate reference standard; test and reference standard assessed independently of each other; blinded reading of results; patient sample included an appropriate spectrum as in clinical practice; reproducibility and interpretation of test results determined; and description of test method sufficient to permit replication. 1 indicates presence and 0 indicates absence.]

Record# Author Year	**Study Quality Examiner(s) Group(s) and Sample Size**	**Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics**	**Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints**	**Findings**
759 Avery 1986	Diagnostic study Study Quality Score (0–6): 2 (100001) Examiner(s): • pneumatic otoscopy: two validated otoscopists • acoustic reflectometry (Acoustic Otoscope [Endeco Medical]): not specified Study Cohort: Children with proven chronic OME treated surgically (2-year follow-up) N=451 subjects, 4147 observations	Time: not specified Place: Not specified Affiliation: Otitis Media Study Center, Santa Rosa Med. Center & The Univ. of Texas Health Sciences Center, San Antonio, TX Inclusion: • proven chronic OME and treated surgically as part of an ongoing clinical trial • ears with intact tympanic membranes Exclusion: None Patient Characteristics: • age range 4–8 years old	Comparisons: 1. Acoustic reflectometry Dx– : ≤ 3 dB Dx+ : > 3 dB Validated pneumatic otoscopy GS– : fluid absent GS+ : fluid present 2. Acoustic reflectometry Dx– : ≤ 4 dB Dx+ : > 4 dB Validated pneumatic otoscopy GS– : fluid absent GS+ : fluid present 3. Acoustic reflectometry Dx– : ≤ 5 dB Dx+ : > 5 dB Validated pneumatic otoscopy GS– : fluid absent GS+ : fluid present	Unit of measure: Observation Comparison 1: sensitivity 77.74% (908/1168) specificity 68.28% (2034/2979) PPV 49.00% (908/1853) NPV 88.67% (2034/2294) accuracy 70.94% (2942/4147) prevalence 28.17% (1168/4147) Comparison 2: sensitivity 65.24% (762/1168) specificity 83.08% (2475/2979) PPV 60.19% (762/1266) NPV 85.91% (2475/2881) accuracy 78.06% (3237/4147) prevalence 28.17% (1168/4147) Comparison 3: sensitivity 50.43% (589/1168) specificity 92.25% (2748/2979) PPV 71.83% (589/820) NPV 82.60% (2748/3327) accuracy 80.47% (3337/4147) prevalence 28.17% (1168/4147)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
766 Babonis 1991	Diagnostic study Study Quality Score (0–6): 5 (111011) Not defined Examiner(s): • tympanometry (MicroTymp [Welch Allen]) and acoustic reflectometry (model 501 [ENT Medical Devices]): one of the authors, specialty not specified • myringotomy: unspecified "surgeon" Study Cohort: Ears in children scheduled for elective myringotomy and PE tube placement N=120 potential subjects and 240 potential ears; 220 ears studied.	Time: 12/1988–11/1989 Place: Not specified Affiliation: Department of Pediatrics, Madigan Army Medical Center, Tacoma, WA Inclusion: • scheduled for elective myringotomy and PE tube placement Exclusion: • poor patient cooperation • mechanical malfunction precluding hard-copy recordings • patient PE tubes • ears microtic, patient undergoing myringotomy and PE tube placement in only 1 ear, and chronic TM perforation Patient Characteristics: • mean age 37.8 months, median 30 months, range 6 months to 10 years 9 months • male 139 ears	Comparisons: 1. Portable tympanometer Dx– : 59–151 daPa Dx+ : >151 daPa Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Portable tympanometer Dx– : A, CI, Ch, As, pA Dx+ : B Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Acoustic reflectometry Dx– : ≤ 5 RU Dx+ : > 5 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 85.59% (101/118) specificity 57.84% (59/102) PPV 70.14% (101/144) NPV 77.63% (59/76) accuracy 72.73% (160/220) prevalence 53.64% (118/220) Comparison 2: sensitivity 77.97% (92/118) specificity 82.35% (84/102) PPV 83.64% (92/110) NPV 76.36% (84/110) accuracy 80.00% (176/220) prevalence 53.64% (118/220) Comparison 3: sensitivity 57.63% (68/118) specificity 88.24% (90/102) PPV 85.00% (68/80) NPV 64.29% (90/140) accuracy 71.82% (158/220) prevalence 53.64% (118/220)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
784 Barnett 1998	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (Race Car Model® [American Electromedics]): research assistant • acoustic reflectometry (EarCheck PRO® [MDI Instruments]): research assistant • myringotomy: not specified *The two models used to measure acoustic reflectometry gave identical results – only reporting the results for EarCheck PRO Study Cohort: Children enrolled in study who had a scheduled surgery for or were receiving placement of tympanostomy tubes that day N=193 subjects enrolled Visit1~N=150 subjects, 274 ears Visit2~N=155 subjects, 299 ears *150 children were examined at visit 1; surgery was performed on 155 children and 299 ears were examined at visit 2.	Time: not specified Place: Children's Hospital, Boston, MA Affiliation: as above Inclusion: • Age: 6 months–14 yrs • scheduling surgery for or receiving placement of tymp. tubes that day Exclusion: • craniofacial abnormality • tympanostomy tubes in place or having had them placed w/in past 6 months • fluid draining from ear Patient Characteristics: • mean age 4.07 years, range 0.51–14.6 years • male 101 • Caucasian 102, Black 25, Asian 3, other 25	Comparisons: 1. Professional tympanometry Dx– : peak compliance > 0.1 Dx+ : peak compliance ≤ 0.1 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry Dx– : peak compliance > 0.2 Dx+ : peak compliance ≤ 0.2 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Acoustic reflectometry Dx– : > 95 degr. (spec.grad. angle) Dx+ : ≤ 95 degrees Myringotomy (sedated) GS– : Fluid absent GS+ : Fluid present 4. Acoustic reflectometry Dx– : ≥ 49 degrees (spec. grad.) Dx+ : < 49 degrees Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 54.29% (95/175) specificity 83.87% (104/124) PPV 82.61% (95/115) NPV 56.52% (104/184) accuracy 66.56% (199/299) prevalence 58.53% (175/299) Comparison 2: sensitivity 62.86% (110/175) specificity 75.00% (93/124) PPV 78.01% (110/141) NPV 58.86% (93/158) accuracy 67.89% (203/299) prevalence 58.53% (175/299) Comparison 3: sensitivity 95.43% (167/175) specificity 31.45% (39/124) PPV 66.27% (167/252) NPV 82.98% (39/47) accuracy 68.90% (206/299) prevalence 58.53% (175/299) Comparison 4: sensitivity 38.29% (67/175) specificity 92.74% (115/124) PPV 88.16% (67/76) NPV 51.57% (115/223) accuracy 60.87% (182/299) prevalence 58.53% (175/299)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
810 Beery 1975	Diagnostic study Study Quality Score (0–6): 2 (100001) Examiner(s): • tympanometry (Grason-Stadler 1720): not specified • myringotomy: not specified Study Cohort: Children with history of recurrent AOM, or evidence otoscopically of persistent MEE, or both N=70 subjects, 129 ears	Time: not specified Place: Not specified Affiliation: Dept. of Otolaryngology, Tufts Univ., MA Inclusion: • Age: 2–15 years • candidate for bilateral myringotomy and insertion of tubes on basis of children's history with concurrent or recent middle ear disease Exclusion: None Patient Characteristics: • mean age 6.7 years, range 2–15 years • male 43, female 27	Comparisons: 1. Professional tympanometry[a] Dx– : See notes section below Dx+ : See notes section below; Class II (at B220) if ≥ 0.3mmh Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 95.65% (66/69) specificity 90.00% (54/60) PPV 91.67% (66/72) NPV 94.74% (54/57) accuracy 93.02% (120/129) prevalence 53.49% (69/129)

[a] DM - pos. cutpoint: For trace at B660 Hz: Class I if suscept. <0.4 mmho, Class III, Class IV; For trace at B220 Hz - Class II if susceptance <0.3 mmho.
DM - negative cutpoint: For trace at B660 Hz - Class I if susc. > or = 0.4 mmh; For trace at B220 Hz - Class II if susceptance ≥0.3 mmho.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
817 Ben-David 1981	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • audiometry (Maico MA-24 and Beltone 15 CX): not specified • tympanometry (Amplaid 702 and American 83): not specified • myringotomy: operating surgeon *Air-bone gap was used in this study as the tested diagnostic method. Study Cohort: Children who underwent tympanometric and audiometric evaluations as well as myringotomy N=157 subjects, 311 ears	Time: not specified Place: Not specified Affiliation: Rothschild Univ. Hospital, Haifa, Israel Inclusion: • Age: 6 months–14 years • evidence of SOM, which failed to ease after conservative Tx with decongestants, antibiotics, and systemic steroids • candidate for myringotomy and tympanostomy-tube insertion Exclusion: None Patient Characteristics: • age range 6 months–14 years	Comparisons: 1. Professional tympanometry[a] Dx– : Type A and Type C Dx+ : Type B Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Audiometry–air and bone conduction thresholds[b] Dx– : < 15 db air-bone gap Dx+ : 15–40 db air-bone gap Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 78.00% (195/250) specificity 63.93% (39/61) PPV 89.86% (195/217) NPV 41.49% (39/94) accuracy 75.24% (234/311) prevalence 80.39% (250/311) Comparison 2: sensitivity 73.91% (153/207) specificity 68.18% (45/66) PPV 87.93% (153/174) NPV 45.45% (45/99) accuracy 72.53% (198/273) prevalence 75.82% (207/273)

[a] We assumed they considered Type B as positive for effusion and all other types negative although they never state this in the text.
[b] The authors never officially define negative and positive cutoff points for the diagnostic method; we assumed that the cutoff points we have supplied above are correct based on their comments in the discussion section.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
886 Block 1998	Diagnostic study Study Quality Score (0–6): 6 (111111) Examiner(s): • tympanometry (American Electromedic RaceCar®): not specified • acoustic reflectometry (EarCheck PRO®): not specified • validated pneumatic otoscopy: validated experienced pediatric otoscopists Study Cohort: Children who participated in study to assess diagnostic accuracy of EarCheck PRO and tympanometry *(Three additional studies were reported in this article but none reported findings relevant to key question four.) N=528 subjects, 870 ears	Time: not specified Place: Validation study recruited from four sites: Barstown, KY; Pittsburgh, PA; Scottsdale, AZ; Rochester,NY/ Long. Study ~ Weston Pediatrics, Weston, MA Affiliation: Kentucky Pediatric Research, Inc.; Children's Hospital of Pittsburgh; Scottsdale Pediatric Center; Elmwood Pediatric Group (NY); Weston Pediatric Physicians; MDI Instruments, Inc. (MA); Capital Management Consulting, Inc. (MA) Inclusion: • Age: 6 months–18 yrs • Exclusion: Myringotomy tubes last 6 months • Perforated tympanic membrane • Cerumen that projected into >50% of the radius of the ear canal • Any serious or psychologic illness • Immunodeficiencies • craniofacial deformities and/or middle ear abnormalities Patient Characteristics: • Age: 6–11 months (N=27); 5 years (N=261); 6–10 years (N=161); 11–18 years (N=79) • Male 51%, Female 49%	Comparisons: 1. Professional tympanometry[a] Dx– : See notes below. Peak-compensated static acoustic admittance 0.22–1.5 mmhos Dx+ : See notes below. Peak acoustic admittance <0.22 mmhos Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 2. Acoustic reflectometry Dx– : ≥70 degrees SG-AR Dx+ : ≤ 69 degrees SG-AR Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present	Unit of measure: Ear Comparison 1: sensitivity 33.92% (58/171) specificity 89.12% (598/671) PPV 44.27% (58/131) NPV 84.11% (598/711) accuracy 77.91% (656/842) prevalence 20.31% (171/842) (Note: PPV and NPV numbers in Table 1 of the article are incorrect.) Comparison 2: sensitivity 67.43% (118/175) specificity 87.05% (605/695) PPV 56.73% (118/208) NPV 91.39% (605/662) accuracy 83.10% (723/870) prevalence 20.11% (175/870)

[a] DM – negative cutoff: equivalent ear canal volume 0.4–2.3 cc; tympanometric peak pressure +200 to –400 daPa; tympanometric Width <180 daPa; DM – positive cutoff: absent tympanometric peak; tympanometric Width >180.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
888 Bluestone 1973	Diagnostic study Study Quality Score (0–6): 2 (100001) Examiner(s): • tympanometry (Madesen ZO-70 Electroacoustic Impedance Bridge): not specified • audiometry (Maico portable audiometer): not specified Study Cohort: Children who were studied to determine the relative validity of air-conduction audiometry and tympanometry in predicting presence of middle ear effusion N=84 subjects Comparison 1: 52 subjects, 87 ears Comparison 2: 55 subjects, 91 ears	Time: not specified Place: Not specified Affiliation: Cleft Palate Center and the Depts. Of Otolaryngology and Pediatrics, Univ. of Pittsburgh, PA; Children's Hospital of Pittsburgh; Dept. of Otolaryngology, Tufts Univ. Boston City Hospital, MA Inclusion: • Age: 2 months–15 years • history of recurrent AOM, or evidence otoscopically of persistent MEE, or both (considered candidates for myringotomy and tympanostomy tube insertion) Exclusion: None Patient Characteristics: • mean age 5.3 years • male 64, female 20	Comparisons: 1. Professional tympanometry[a] Dx– : Patterns 1,2, and 3 Dx+ : Patterns 4 and 5 Myringotomy (sedated) GS– : Effusion absent GS+ : Effusion present 2. Audiometry – air conduction threshold Dx– : Below 25 dB Dx+ : ≥ 25 dB Myringotomy (sedated) GS– : Effusion absent GS+ : Effusion present 3. Audiometry – air and bone conduction thresholds Dx– : Not Defined Dx+ : Not Defined Myringotomy (sedated) GS– : Effusion absent GS+ : Effusion present	Unit of measure: Ear Comparison 1: sensitivity 100.00% (58/58) specificity 55.17% (16/29) PPV 81.69% (58/71) NPV 100.00% (16/16) accuracy 85.06% (74/87) prevalence 66.67% (58/87) Comparison 2: sensitivity 51.72% (30/58) specificity 75.76% (25/33) PPV 78.95% (30/38) NPV 47.17% (25/53) accuracy 60.44% (55/91) prevalence 63.74% (58/91) Comparison 3[b]: Not done

[a] Pattern 1 = Normal Compliance, Normal Pressure; 2 = Nml comp, negative pressure; 3 = High comp, nml pressure; 4 = low comp, nml pressure; 5 = low comp, pressure high negative, positive, or indeterminate.
[b] Could not be completed because negative and positive cut-points are not defined for Air-Bone Gap.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
889 Bluestone 1979	Diagnostic study Study Quality Score (0–6): 4 (110011) Examiner(s): • otoscopy: two otolaryngologists, two pediatricians • tympanometry and middle ear muscle reflex (Model Z073 [Madsen Electronics]): by or under the direction of an audiologist, classified independently by two audiologists • myringotomy: surgeon Study Cohort: Validity of diagnosing OME by otoscopy, tympanometry, and middle ear muscle reflex ~ gold standard: Myringotomy (referenced as Study 2 in article) *(an additional study was reported in this article but findings were not relevant to key question four) N=239 subjects, 425 ears	Time: S2 ~ 9/1/1977–2/22/1978 Place: not specified Affiliation: Dept. of Otolaryngology, Children's Hospital of Pittsburgh, PA Inclusion: • Age: 7 months–15 years • patients scheduled for bilateral myringotomy and insertion of tympanostomy tubes on basis of a history of recurrent acute and/or persistent OME Exclusion: None Patient Characteristics: As above	Comparisons: 1. Pneumatic otoscopy – examiner Validation not specified Dx– : OME absent (– cases) Dx+ : OME present (+ and +/– cases) Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry[a] Dx– : Paradise classification variants 1–6, 9,15 Dx+ : Paradise classification variants 7,8,10–14 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Examiner type (Otolaryngologist) Pneumatic otoscopy – unvalidated examiner Dx– : OME absent (– cases) Dx+ : OME present (+ and +/– cases) Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 94.53% (242/256) specificity 77.51% (131/169) PPV 86.43% (242/280) NPV 90.34% (131/145) accuracy 87.76% (373/425) prevalence 60.24% (256/425) Comparison 2: sensitivity 93.75% (240/256) specificity 70.41% (119/169) PPV 82.76% (240/290) NPV 88.15% (119/135) accuracy 84.47% (359/425) prevalence 60.24% (256/425) Comparison 3: sensitivity 100.00% (29/29) specificity 22.73% (5/22) PPV 63.04% (29/46) NPV 100.00% (5/5) accuracy 66.67% (34/51) prevalence 56.86% (29/51) *Continued on next page*

[a] Middle ear muscle reflex testing was also evaluated as a diag. method in this study, but wasn't included in the list of diagnostic test choices.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 889 Bluestone 1979			*Continued* 4. Examiner type (Pediatrician) Pneumatic otoscopy – unvalidated examiner Dx– : OME absent (– cases) Dx+ : OME present (+ and +/– cases) Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 5. Examiner type (Pediatrician) Pneumatic otoscopy – unvalidated examiner Dx– : OME absent (– cases) Dx+ : OME present (+ and +/– cases) Myringotomy (sedated) GS– : fluid absent GS+ : fluid present)	*Continued* Unit of measure: Ear Comparison 4: sensitivity 96.30% (26/27) specificity 25.00% (5/20) PPV 63.41% (26/41) NPV 83.33% (5/6) accuracy 65.96% (31/47) prevalence 57.45% (27/47) Comparison 5: sensitivity 76.00% (19/25) specificity 50.00% (7/14) PPV 73.08% (19/26) NPV 53.85% (7/13) accuracy 66.67% (26/39) prevalence 64.10% (25/39)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
989 Cantekin 1977	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • tympanometry (types described under each comparison): not specified • myringotomy: not specified Study Cohort: Children scheduled for bilateral myringotomy and insertion of tympanostomy tubes for recurrent acute otitis media history or persistent middle ear effusion by otoscopy N=68 subjects, 120 ears	Time: not specified Place: Not specified Affiliation: Dept. of Otolaryngology, Children's Hospital of Pittsburgh and the Univ. of Pittsburgh, PA Inclusion: • Age: 18 months–14 years • history of recurrent AOM or otoscopic evidence of persistent MEE, or both • scheduled for bilateral myringotomy and insertion of tympanostomy tubes for recurrent acute otitis media history or persistent middle ear effusion by otoscopy Exclusion: None Patient Characteristics: • median age 6.5 years • male 43, female 25	Comparisons: 1. Professional tympanometry[a] (Madsen Electro-Acoustic Impedance Bridge, model ZO-70) Dx– : air pressure ≥ –100 mm H2O; compliance <5.0 OR (See notes.) Dx+ : air press <–100 mm H_2O; compliance ≥ 5 OR (See notes.) Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry[b] (Grason-Stadler Otoadmittance meter, model 1720) Dx– : See notes below. Dx+ : See notes below. Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 84.13% (53/63) specificity 87.72% (50/57) PPV 88.33% (53/60) NPV 83.33% (50/60) accuracy 85.83% (103/120) prevalence 52.50% (63/120) Comparison 2: sensitivity 84.13% (53/63) specificity 87.72% (50/57) PPV 88.33% (53/60) NPV 83.33% (50/60) accuracy 85.83% (103/120) prevalence 52.50% (63/120)

[a] Negative cutoff point (continued): air pressure <–100 mm H2O; compliance <5.0 Madsen units; gradient ≥ 0.5 Madsen units. Positive Cut Point (continued): air pressure <–100 mm H2O; compliance ≥5.0 Madsen units; gradient <0.5 Madsen units.

[b] Negative cutoff: Class I curve with peak susceptance > or = .4mmho at 660 Hz; Class II curve with peak ≥.3 mmhos at 220 Hz. Postive cutoff: Class I curve, peak < .4mmho at 660 Hz; Class II curve, peak <.3mmho at 220 Hz; Class III and IV.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
990 Cantekin 1980	Diagnostic study Study Quality Score (0–6): 4 (110011) Examiner(s): • otoscopy: two otolaryngologists • tympanometry and middle ear muscle reflex (Model ZO73 [Madsen Electronics]): by or under the direction of an audiologist, independently classified by two investigators • myringotomy: surgeon Study Cohort: Children who underwent myringotomy ~ myringotomy findings were compared to OME diagnosis based on otoscopy, tympanometry and ME muscle reflex N=333 subjects, 599 ears	Time: not specified Place: Not specified Affiliation: Dept. of Otolaryngology, Univ. of Pittsburgh School of Medicine, and Children's Hospital of Pittsburgh, PA Inclusion: • Age: 7 months–15 years • patients scheduled for bilateral myringotomy and insertion of tympanostomy tubes • history of recurrent AOM, or otoscopic evidence of persistent OME, or both Exclusion: None Patient Characteristics: • median age 6 years • male 203, female 130	Comparisons: 1. Pneumatic otoscopy[a] – unvalidated examiner Dx– : OME absent Dx+ : OME present Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Pneumatic otoscopy[b] – unvalidated examiner Dx– : OME absent Dx+ : OME present Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Quantitative tympanometry[c] Dx– : ME muscle reflex threshold ≤105dB Dx+ : Absence of ME muscle reflex Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 4. Quantitative tympanometry[d] Dx– : ME muscle reflex threshold ≤105dB Dx+ : Absence of ME muscle reflex Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 96.96% (223/230) specificity 81.43% (114/140) PPV 89.56% (223/249) NPV 94.21% (114/121) accuracy 91.08% (337/370) prevalence 62.16% (230/370) Comparison 2: sensitivity 87.56% (176/201) specificity 80.54% (120/149) PPV 85.85% (176/205) NPV 82.76% (120/145) accuracy 84.57% (296/350) prevalence 57.43% (201/350) Comparison 3: sensitivity 96.30% specificity 34.60% PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 4: sensitivity 92.90% specificity 52.00% PPV unknown NPV unknown accuracy unknown prevalence unknown *Continued on next page*

[a] These are results for otolaryngologist A.
[b] These results are for otolaryngologist B.
[c] ME muscle reflex measured using ipsilateral stimulation and ambient pressure; Stimulus frequency = 1000.
[d] ME muscle reflex measured using ipsilateral stimulation and tympanogram's peak pressure; Stimulus frequency = 1000.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 990 Cantekin 1980			*Continued* 5. Quantitative tympanometry[a] Dx– : ME muscle reflex threshold ≤105dB Dx+ : Absence of ME muscle reflex Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 6. Quantitative tympanometry[b] Dx– : ME muscle reflex threshold ≤105dB Dx+ : Absence of ME muscle reflex Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 7. Quantitative tympanometry[c] Dx– : ME muscle reflex threshold ≤105dB Dx+ : Absence of ME muscle reflex Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 5: sensitivity 97.90% specificity 24.80% PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 6: sensitivity 95.50% specificity 37.40% PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 7: sensitivity 95.50% specificity 37.40% PPV unknown NPV unknown accuracy unknown prevalence unknown *Continued on next page*

[a] ME muscle reflex measured using contralateral stimulation and ambient pressure; Stimulus frequency = 1000.
[b] ME muscle reflex measured using contralateral stimulation and tympanogram's peak pressure; Stimulus frequency = 1000.
[c] ME muscle reflex measured using ipsilateral stimulation and ambient pressure; Stimulus frequency = 2000.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 990 Cantekin 1980			*Continued* 8. Quantitative tympanometry[a] Dx– : ME muscle reflex threshold ≤105dB Dx+ : Absence of ME muscle reflex Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 9. Quantitative tympanometry[b] Dx– : ME muscle reflex threshold ≤105dB Dx+ : Absence of ME muscle reflex Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 10. Quantitative tympanometry[c] Dx– : ME muscle reflex threshold ≤105dB Dx+ : Absence of ME muscle reflex Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear Comparison 8: sensitivity 91.90% specificity 50.70% PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 9: sensitivity 97.10% specificity 29.40% PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 10: sensitivity 93.30% specificity 39.40% PPV unknown NPV unknown accuracy unknown prevalence unknown

[a] ME muscle reflex measured using ipsilateral stimulation and tympanogram's peak pressure; Stimulus frequency = 2000.
[b] ME muscle reflex measured using contralateral stimulation and ambient pressure; Stimulus frequency = 2000.
[c] ME muscle reflex measured using contralateral stimulation and tympanogram's peak pressure; Stimulus frequency = 2000.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1238 Fiellau-Nikolajsen 1980	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (Madsen Impedance Meter, type ZO-73, testing tone 220Hz): the author • myringotomy: two otosurgeons *Same study population as record #1241 Study Cohort: Children whose ear(s) persistently showed abnormal screening-tympanometry over 6 months; myringotomic aspiration of any MEE was performed N=44 subjects (21 children bilateral, 23 unilateral), 88 ears	Time: Intitial screening in Aug. 1978; operation (if needed) in Feb. 1979 Place: Subjects identified from tympanometric screening in a Danish provincial municipality Affiliation: ENT Dept. Hjoerring Hospital, Denmark Inclusion: • Age: 3 years at initial screening • one or both ears consistently showed abnormal screening-tympanometry in Aug., Sept., Nov. 1978 and Feb. 1979 • living in municipality of Hjoerring, Denmark Exclusion: None Patient Characteristics: • mean age 47.2 months, range 42–54 months • male 23, female 21	Comparisons : 1.1 Professional tympanometry Dx– : Compliance value >.10 ccm Dx+ : Compliance value ≤.10 ccm Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 1.2 Professional tympanometry Dx– : Compliance value >.20 ccm Dx+ : Compliance value ≤.20 ccm Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 1.3 Professional tympanometry Dx– : Compliance value >.30 ccm Dx+ : Compliance value ≤.30 ccm Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1.1[a]: sensitivity 19.57% (9/46) specificity 100.00% (42/42) PPV 100.00% (9/9) NPV 53.16% (42/79) accuracy 57.95% (51/88) prevalence 52.27% (46/88) Comparison 1.2 [a]: sensitivity 45.65% (21/46) specificity 95.24% (40/42) PPV 91.30% (21/23) NPV 61.54% (40/65) Accuracy 69.32% (61/88) prevalence 52.27% (46/88) Comparison 1.3 [a]: sensitivity 65.22% (30/46) specificity 76.19% (32/42) PPV 75.00% (30/40) NPV 66.67% (32/48) Accuracy 70.45% (62/88) prevalence 52.27% (46/88) *Continued on next page*

[a] Estimated from figure 3.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1238 Fiellau-Nikolajsen 1980			*Continued* 2. Professional tympanometry Dx– : Gradient > .05 ccm Dx+ : Gradient ≤ .05 ccm Myringotomy (sedated) GS– : Fluid absent GS+ : Fluid present 3. Professional tympanometry Dx– : Gradient > .1 ccm Dx+ : Gradient ≤ .1 ccm Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear Comparison 2[a]: sensitivity 78.26% (36/46) specificity 92.86% (39/42) PPV 92.31% (36/39) NPV 79.59% (39/49) accuracy 85.23% (75/88) Comparison 3 [a]: sensitivity 91.30% (42/46) specificity 54.76% (23/42) PPV 68.85% (42/61) NPV 85.19% (23/27) accuracy 73.86% (65/88) prevalence 52.27% (46/88)

[a] These values were calculated from percentages that were provided in the text and thus their accuracy is uncertain.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1241 Fiellau-Nikolajsen 1980	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (Madsen Impedance Meter, type ZO-73, testing tone 220Hz): the author • myringotomies: two otosurgeons * This is the same study as 1238. However, the data are presented slightly differently. Study Cohort: Children who showed abnormal tympanometry and underwent myringotomy. N=44 subjects, 88 ears	Time: 8/1978–2/1979 (tested in August, 1978 – retests in September and November, 1978 and February,1979) Place: Danish urban municipality (Hjoerring munipality: pop. 37,500) Affiliation: Dept's of Otolarygology and Audiology, Hjoerring Hospital, Hjoerring, Denmark Inclusion: • Age: 3 years at study entry (42–54 months by Feb. 1979) • abnormal tympanometric results persisting through 6 months Exclusion: None Patient Characteristics: • mean age 47.2 months, range 42–54 months • male 23, female 21	Comparisons: 1. Professional tympanometry Dx– : Type A and C1 tympanogram Dx+ : Type B and C2 tympanograms Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1[a]: sensitivity 91.30% (42/46) specificity 88.10% (37/42) PPV 89.36% (42/47) NPV 90.24% (37/41) accuracy 89.77% (79/88) prevalence 52.27% (46/88)

[a] Numbers were calculated from percentages and total "n's" presented in Table V.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1245 Fiellau-Nikolajsen 1983	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (Madsen tympanometer, model ZO-73, [Madsen Electronics], 220 Hz testing tone): not specified • myringotomy: ENT surgeon *This is the same study as 1238. However, the data are presented slightly different. Study Cohort: Children who underwent Myringotomy due to persistent abnormal tympanograms across four screens *(Grp 1 was derived from an overall sample of 404 subjects) N=44 subjects, 88 ears	Time: 1978–1979 (four screens: Aug., Sep. & Nov., 1978, Feb. 1979) Place: Subjects identified in municipality of Hjoerring, Denmark, (medium-size urban town with a population of approximately 20,000) Affiliation: ENT Dept., Hjoerring Hospital, Denmark Inclusion: • Age: 3 years (36–48 months) • living in municipality of Hjoerring, Denmark Exclusion: None Patient Characteristics: • age range 36–48 months	Comparisons: 1. Professional tympanometry Dx– : Middle ear pressure >–100 mm H2O or non-flat curve Dx+ : Middle ear pressure ≤–100 mm H2O or flat curve Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry Dx– : Middle ear pressure >–200 mm H2O or non-flat curve Dx+ : Middle ear pressure ≤–200 mm H2O or flat curve Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Professional tympanometry Dx– : Non-flat curve Dx+ : Flat curve Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 95.65% (44/46) specificity 64.29% (27/42) PPV 74.58% (44/59) NPV 93.10% (27/29) accuracy 80.68% (71/88) prevalence 52.27% (46/88) Comparison 2: sensitivity 91.30% (42/46) specificity 88.10% (37/42) PPV 89.36% (42/47) NPV 90.24% (37/41) accuracy 89.77% (79/88) prevalence 52.27% (46/88) Comparison 3: sensitivity 82.61% (38/46) specificity 100.00% (42/42) PPV 100.00% (38/38) NPV 84.00% (42/50) accuracy 90.91% (80/88) prevalence 52.27% (46/88)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1250 Finitzo 1992	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (Maico Screening Immittance Bridge, model No. 610): certified audiologist • pneumatic otoscopy: pediatric otolaryngologist • myringotomy: not specified *No information regarding when tympanometry occurred relative to myringotomy is provided in the methods section. Study Cohort: Children who were to undergo myringotomy for the placement of ventilation tubes for recurrent OM or persistent OME N=88 subjects, 163 ears	Time: not specified Place: Children's Medical Center in Dallas, Texas Affiliation: Methodist Medical Center, Dallas, TX; Callier Center for Communication Disorders and Dept. of Otolaryngology, Univ. of Texas, Dallas, TX Inclusion: • Age: 6 months–9 years • undergoing myringotomy for tube placement to treat recurrent OM or persisitent OME Exclusion: None Patient Characteristics: • mean age 31 months • male 59, female 27	Comparisons: 1. Pneumatic otoscopy[a] – unvalidated examiner Dx– : effusion absent Dx+ : effusion present Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry[a] Dx– : Type A or A' tympanogram Dx+ : Type B tympanogram Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 93.04% (107/115) specificity 58.33% (28/48) PPV 84.25% (107/127) NPV 77.78% (28/36) accuracy 82.82% (135/163) prevalence 70.55% (115/163) Comparison 2: sensitivity 90.28% (65/72) specificity 86.36% (19/22) PPV 98.59% (65/68) NPV 73.08% (19/26) accuracy 89.36% (84/94) prevalence 76.60% (72/94)

[a] Of the 94 ears included the majority underwent myringotomy for recurrent AOM rather than OME; results are not stratified according to rationale for myringotomy. Of 86 children, only 16 had persistent OME; the rest had recurrent OM.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1280 Freyss 1980	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • tympanometry (Madsen ZO 73 impedance bridge): not specified • myringotomy: same investigator for all children Study Cohort: Children undergoing bilateral myringtomy for recurrent otitis media or respiratory tract infection N=50 subjects, 99 ears	Time: not specified Place: Not specified Affiliation: Univ. of Paris VII, Dept. of Oto-laryngology, Hospital Lariboisiere; Dept. of Oto-laryngology, Children's Hosptial Bretonneau, Paris, France Inclusion: • Age: 6 months–8 years • undergoing bilateral myringtomy for recurrent otitis media or respiratory tract infection Exclusion: None Patient Characteristics: • age range 6 months–8 years	Comparisons: 1. Quantitative tympanometry Dx– : ≤115 dB (at 1000Hz) Dx+ : > 115 dB (at 1000Hz) Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Quantitative tympanometry Dx– : < 2 STAR (suprathreshold acoustic reflex) (Madsen units) Dx+ : ≥2 STAR (Madsen units) Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Quantitative tympanometry Dx– : ≤-100 mmH2O Peak Pressure Dx+ : >-100 mmH2O Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 75.00% (57/76) specificity 82.61% (19/23) PPV 93.44% (57/61) NPV 50.00% (19/38) accuracy 76.77% (76/99) prevalence 76.77% (76/99) Comparison 2: sensitivity 93.42% (71/76) specificity 65.22% (15/23) PPV 89.87% (71/79) NPV 75.00% (15/20) accuracy 86.87% (86/99) prevalence 76.77% (76/99) Comparison 3: sensitivity 52.63% (40/76) specificity 82.61% (19/23) PPV 90.91% (40/44) NPV 34.55% (19/55) accuracy 59.60% (59/99) prevalence 76.77% (76/99) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1280 Freyss 1980			*Continued* 4. Quantitative tympanometry Dx– : ≤ 4 Madsen units Maximum Compliance Dx+ : > 4 Madsen Units Maximum Compliance Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 5. Quantitative tympanometry Dx– : <1 Madsen Unit Tympanometric Gradient Dx+ : ≥1 Madsen Unit Tympanometric Gradient Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear Comparison 4: sensitivity 51.32% (39/76) specificity 78.26% (18/23) PPV 88.64% (39/44) NPV 32.73% (18/55) accuracy 57.58% (57/99) prevalence 76.77% (76/99) Comparison 5: sensitivity 61.84% (47/76) specificity 78.26% (18/23) PPV 90.38% (47/52) NPV 38.30% (18/47) accuracy 65.66% (65/99) prevalence 76.77% (76/99)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1282 Fria 1980	Diagnostic study Study Quality Score (0–6): 3 (100101) Examiner(s): • tympanometry (Grason-Stadler, model 1722, 1000Hz at 102 dB SPL): not specified • myringotomy: not specified Study Cohort: OME ~ Children with history of recurrent AOM or otoscopic evidence of persistent OME, or both *(N=40 otoscopically normal children were reported in this article but findings were not relevant to key question four) N=172 subjects, 344 ears	Time: not specified Place: Normal children were from the ambulatory care center. Affiliation: Dept. of Otolaryngology, Univ. of Pittsburgh School of Medicine; Audiology Division, Dept. of Otolaryngology, Children's Hospital of Pittsburgh, Pittsburgh, PA Inclusion: • Age: 3–15 years • history of recurrent AOM or otoscopic evidence of persistent OME, or both Exclusion: None Patient Characteristics: • Grp 1 mean age 8 years, range 3–15 years, Grp 2 mean age 4 years, range 7 months–13 years	Comparisons:[a] 1. Portable tympanometer Dx– : Middle Ear Analyzer Regions 1,2, and 3 Dx+ : MEA region 4 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Portable tympanometer Dx– : acoustic reflex present Dx+ : acoustic reflex absent Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Portable tympanometer Dx– : Pressure Peak Present Dx+ : Pressure Peak Absent Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 96.50% (138/143) specificity 62.28% (71/114) PPV 76.24% (138/181) NPV 93.42% (71/76) accuracy 81.32% (209/257) prevalence 55.64% (143/257) Comparison 2: sensitivity 98.60% (141/143) specificity 44.74% (51/114) PPV 69.12% (141/204) NPV 96.23% (51/53) accuracy 74.71% (192/257) prevalence 55.64% (143/257) Comparison 3: sensitivity 86.71% (124/143) specificity 80.70% (92/114) PPV 84.93% (124/146) NPV 82.88% (92/111) accuracy 84.04% (216/257) prevalence 55.64% (143/257)

[a] Combines children undergoing myringotomy for recurrent AOM and persistent OME. Groups are not separated in the analyses.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
4879 Fried 1985	Diagnostic Study Study Quality Score (0–6): 1 (100000) Examiner(s): • pneumatic otoscopy: three surgeons • acoustic reflectometry (Acoustic Otoscope): not specified Study Cohort: Children undergoing bilateral myringotomy and ventilation tube insertion N=59 subjects, 118 ears	Time: not specified Place: Not specified Affiliation: Joint Center for Otolaryngology, Harvard Medical School, Boston, MA Inclusion: • Age: 1–13 years • undergoing bilateral myringotomy and ventilation tube insertion Exclusion: None Patient Characteristics: • age range 1–13 years	Comparisons: 1. Acoustic reflectometry Dx– : ≤ 4 Reflectivity Units Dx+ : > 4 Reflectivity Units Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1:[a] outcome~(ears – defined by group) sensitivity 70.97% (44/62) specificity 90.00% (36/40) PPV 91.67% (44/48) NPV 66.67% (36/54) accuracy 78.43% (80/102) prevalence 60.78% (62/102)

[a] The results table includes 102 ears. However, this study was of 118 ears. The authors do not discuss why the other 16 ears were excluded from the results table (Table 1 in the article).

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
4878 Gersdorff 1986	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • tympanometry (Grason-Stadler 1723 oto-admittance meter, 220Hz tone): not specified • myringotomy: ENT surgeon Study Cohort: Pathological children with serous and secretory otitis media who underwent myringotomy and ventilating tube insertion (50% of these children received denitrogenization before induction of classic inhalation anesthesia) *(A control group was studied; however, findings were not relevant to key question four.) N=64 subjects, 128 ears	Time: not specified Place: Not specified Affiliation: Dept. of Otorhinolaryngology, Dept. of Anesthesiology, University of Louvain, Belgium Inclusion: • classified as pathological due to presence of serous and secretory OM upon clinical examination Exclusion: None Patient Characteristics: • combined mean age 7.5 years	Comparisons:[a] 1. Professional tympanometry Dx– : Type I tympanogram Dx+ : Type IV tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry Dx– : Typanogram types I, II, III, & V Dx+ : Typanogram type IV Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Professional tympanometry Dx– : Tympanogram Type I Dx+ : Tympanogram types II,III,IV,&V Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 89.71% (61/68) specificity 28.57% (8/28) PPV 75.31% (61/81) NPV 53.33% (8/15) accuracy 71.88% (69/96) prevalence 70.83% (68/96) Comparison 2: sensitivity 75.31% (61/81) specificity 57.45% (27/47) PPV 75.31% (61/81) NPV 57.45% (27/47) accuracy 68.75% (88/128) prevalence 63.28% (81/128) Comparison 3: sensitivity 91.36% (74/81) specificity 17.02% (8/47) PPV 65/49% (74/113) NPV 53.33% (8/15) accuracy 64.06% (82/128) prevalence 63.28% (81/128)

[a] These data are for tympanograms performed before induction of anesthesia only. Tympanogram classifications: Type I~normal; Type II~endotympanic depression; Type III~endotympanic depression and tympanometric curve rounded by weaker compliance; Type IV~flattened curve; Type V~normal endotympanic pressure and rounded tympanometric curve.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1384 Grimaldi 1976	Diagnostic study Study Quality Score (0–6): 2 (100001) Examiner(s): • audiometry (Peters AP 5 audiometer): OME audiologist • tympanometry (Peters AP 61 impedance meter): one audiologist • binocular micro-tympanoscopy: not specified • myringotomy: single surgeon • visual inspection: effusion probably or possible = OME *Although tympanometry and air-conduction audiometry were also assessed in this study, the data cannot be used because they were combined into a diagnostic algorithm rather than being evaluated as individual methods. Study Cohort: Children referred by otologists for presumptive middle ear effusions N=120 subjects, 209 ears	Time: 9-month period (actual dates not specified) Place: Patients from ENT Dept. of King's College Hospital and referred by otologists based on data from a routine outpatient clinic Affiliation: Not specified apart from author being from London Inclusion: • referred by otologists for presumptive middle ear effusions Exclusion: • inadequate audiometric data Patient Characteristics: • age not specified, other than being children	Comparisons: 1. Non-pneumatic otoscopy[a] Dx– : Effusion absent (includes n=47 undecided cases) Dx+ : Effusion present Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Non-pneumatic otoscopy[b] Dx– : Effusion absent Dx+ : Effusion present (includes n=47 undecided cases) Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Binocular micro-tympanoscopy[c] Dx– : no effusion visualized, undecided group Dx+ : effusion visualized Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 85.71% (132/154) specificity 87.27% (48/55) PPV 94.96% (132/139) NPV 68.57% (48/70) accuracy 86.12% (180/209) prevalence 73.68% (154/209) Comparison 2: sensitivity 98.05% (151/154) specificity 36.36% (20/55) PPV 81.18% (151/186) NPV 86.96% (20/23) accuracy 81.82% (171/209) prevalence 73.68% (154/209) Comparison 3: sensitivity 85.71% (132/154) specificity 87.27% (48/55) PPV 94.96% (132/139) NPV 68.57% (48/70) accuracy 86.12% (180/209) prevalence 73.68% (154/209) *Continued on next page*

[a] There were 47 cases where the examiner was undecided; these have been counted as effusion absent (test negative).
[b] There were 47 cases where the examiner was undecided; these have been counted as effusion present (test positive).
[c] This result includes the "undecided" group as normal.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1384 Grimaldi 1976			*Continued* 4. Binocular micro-tympanoscopy[a] Dx– : no effusion visualized Dx+ : effusion visualized; undecided group Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 5. Binocular micro-tympanoscopy[b] Dx– : No effusion visualized Dx+ : Effusion visualized Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear Comparison 4: sensitivity 98.05% (151/154) specificity 36.36% (20/55) PPV 81.18% (151/186) NPV 86.96% (20/23) accuracy 81.82% (171/209) prevalence 73.68% (154/209) Comparison 5: sensitivity 97.78% (132/135) specificity 74.07% (20/27) PPV 94.96% (132/139) NPV 86.96% (20/23) accuracy 93.83% (152/162) prevalence 83.33% (135/162)

[a] The results table includes the "undecided" group as abnormal.
[b] There were 47 ears that were not classified as "yes" or "no" pre-operatively; rather they were labeled "undecided." These 47 are not included in the results table.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1397 Haapaniemi 1997	Diagnostic study Study Quality Score (0–6): 2 (100100) Examiner(s): • pure-tone audiometry (Madsen TBN 80 audiometer [Madsen Electronics]): not specified • tympanometry and stapedius reflexes (GSI 28 tympanometer [Grason-Stadler], 226 Hz tone): not specified • otomicroscopy: not specified • paracentesis: not specified Study Cohort: 1st and 4th grade children who underwent a routine ENT investigation at school that included: otomicroscopy, pure-tone audiometry, tympanometry and stapedius reflex measurements * (8th grade children were also studied but their age range was above that specified for key question four) N=426 subjects, 850 ears	Time: not specified Place: Elementary schools of Salo, a small town in Finland Affiliation: Dept. Otolaryngology, Univ. Central Hospital of Turku, Finland Inclusion: • Age represented by grade level: First, fourth and eighth grade • attending an elementary school in Salo, Finland Exclusion: None Patient Characteristics: • 1st grade mean age 7.0 years, range 6–9 years; 4th grade mean age 10.2 years, range 10–12 years	Comparisons: 1. Professional tympanometry Dx– : Tympanometric peak pressure (TPP) ≥ –100 daPa Dx+ : TPP < –100 daPa Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry Dx– : Tympanometric peak pressure (TPP) ≥ –100 daPa Dx+: TPP < –100 daPa Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 3. Professional tympanometry Dx– : Tympanometric peak pressure (TPP) ≥ –150 daPa Dx+ : TPP < –150 daPa Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Values for the “b” and “d” cells in the 2X2 table are not provided in article Unit of measure: Varies Comparison 1: Unit of measure: Subjects sensitivity 93.33% (28/30) specificity 88.00% PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 2: Unit of measure: Ear sensitivity 94.74% (36/38) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 3: Unit of measure: Subjects sensitivity 93.33% (28/30) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1397 Haapaniemi 1997			*Continued* 4. Professional tympanometry Dx– : Tympanometric peak pressure TPP≥ –150 daPa Dx+ : TPP < –150 daPa Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 5. Professional tympanometry Dx– : Tympanometric peak pressure TPP ≥ –200 daPa Dx+ : TPP < –200 daPa Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 6. Professional tympanometry Dx– : Tympanometric peak pressure TPP ≥ –200 daPa Dx+: TPP < –200 daPa Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 7. Professional tympanometry Dx– : non-Type B tympanogram Dx+ : Type B tympanogram Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Comparison 4: Unit of measure: Ear sensitivity 94.74%(36/38) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 5: Unit of measure: Subjects sensitivity 86.67% (26/30) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 6: Unit of measure: Ear sensitivity 89.47%(34/38) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 7: Unit of measure: Subjects sensitivity 53.33% (16/30) specificity 100.00% PPV unknown NPV unknown accuracy unknown prevalence unknown *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1397 Haapaniemi 1997			*Continued* 8. Professional tympanometry Dx– : non-Type B tympanogram Dx+ : Type B tympanogram Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 9. Professional tympanometry[a] Dx– : Cutoff of 0.3 ml is given for admittance (see notes below) Dx+ : Cutoff of 0.3 ml is given for admittance (see notes below) Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 10. Professional tympanometry[a] Dx– : Cutoff of 0.3 ml is given for admittance (see notes below) Dx+ : Cutoff of 0.3 ml is given for Admittance (see notes below) Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Comparison 8: Unit of measure: Ear sensitivity 55.26% (21/38) specificity 100.00% PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 9: Unit of measure: Subjects sensitivity 80.00% (24/30) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 10: Unit of measure: Ear sensitivity 81.58% (31/38) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown *Continued on next page*

[a] For diagnostic cutoffs, an Admittance value of 0.3 ml was given as the cutpoint; however, not sure whether admittance < or = 0.3 ml would be a negative result for MEE OR values > 0.3.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1397 Haapaniemi 1997			*Continued* 11. Professional tympanometry[a] Dx– : Relative gradient > 0.35 Dx+ : Relative gradient ≤ 0.35 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 12. Professional tympanometry[a] Dx– : Relative gradient > 0.35 Dx+ : Relative gradient ≤ 0.35 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 13. Professional tympanometry Dx– : Ipsilateral Stapedius Reflex (at 1000Hz) at 105 dB – present Dx+ : Ipsilateral Stapedius Reflex (at 1000Hz) at 105 dB – absent Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 14. Professional tympanometry Dx– : Ipsilateral Stapedius Reflex (at 1000Hz) at 105 dB – present Dx+ : Ipsilateral Stapedius Reflex (at 1000Hz) at 105 dB – absent Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Comparison 11: Unit of measure: Subject sensitivity 76.67% (23/30) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 12: Unit of measure: Ear sensitivity 73.68% (28/38) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 13: Unit of measure: Subject sensitivity 76.67% (23/30) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 14: Unit of measure: Ear sensitivity 78.95% (30/38) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown *Continued on next page*

[a] For diagnostic cutoffs, a Relative Gradient value of 0.35 was given as the cutpoint; however, not sure whether gradient > 0.35 would be a negative result for MEE (I am assuming it is).

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1397 Haapaniemi 1997			*Continued* 15. Professional tympanometry Dx– : Contralateral Stapedius Reflex (at 1000 Hz) 110dB – present Dx+ : Contralateral Stapedius Reflex (at 1000Hz) at 110 dB – absent Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 16. Professional tympanometry Dx– : Contralateral Stapedius Reflex (at 1000 Hz) 110dB – present Dx+ : Contralateral Stapedius Reflex (at 1000Hz) at 110 dB – absent Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 17. Professional tympanometry[a] Dx– : Tympanogram Types A, C1, and C2 Dx+ : Tympanogram Type B Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Comparison 15: Unit of measure: Subject sensitivity 73.33% (22/30) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 16: Unit of measure: Ear sensitivity 73.68% (28/38) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 17: Unit of measure: Ear sensitivity 55.26% (21/38) specificity 100.00% PPV unknown NPV unknown accuracy unknown prevalence unknown *Continued on next page*

[a] These results examine sensitivity of type B tympanograms in detecting OME in children with either transient or prolonged effusion. Prolonged = effusion still present 4 weeks after myringotomy. Transient = ear healthy 4 weeks after myringotomy.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1397 Haapaniemi 1997			*Continued* 18. Audiometry – air conduction threshold Dx– : ≤ 15 hearing threshold at .25 to 8 kHz Dx+ : > 15 dB hearing threshold (at .25 to 8 kHz) Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 19. Audiometry – air conduction threshold Dx– : ≤ 15 hearing threshold at .25 to 8 kHz Dx+ : > 15 dB hearing threshold (at .25 to 8 kHz) Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 20. Audiometry – air conduction threshold Dx– : ≤ 20 dB hearing threshold (.25 to 8 kHz) Dx+ : > 20 dB hearing threshold (.25 to 8 kHz) Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 21. Audiometry – air conduction threshold Dx– : ≤ 20 dB hearing threshold (.25 to 8 kHz) Dx+ : > 20 dB hearing threshold (.25 to 8 kHz) Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present	*Continued* Comparison 18: Unit of measure: Subject sensitivity 82.76% (24/29) specificity 81.36% (323/397) PPV 24.49% (24/98) NPV 98.48% (323/328) accuracy 81.46% (347/426) prevalence 6.81% (29/426) Comparison 19: Unit of measure: Ear sensitivity 86.49% (32/37) specificity 87.00% PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 20: Unit of measure: Subject sensitivity 68.97% (20/29) specificity 87.91% (349/397) PPV 29.41% (20/68) NPV 97.49% (349/358) accuracy 86.62% (369/426) prevalence 8.38% (29/426) Comparison 21: Unit of measure: Ear sensitivity 72.97% (27/37) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence unknown

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1446 Haughton 1977	Diagnostic study Study Quality Score (0–6): 2 (100001) Examiner(s): • tympanometry (instrument not specified): not specified • tympanocentesis: one of several surgeons Study Cohort: S1~ Ears examined in children admitted for aspiration of the middle ear; Haughton study *(Two additional studies were reported in this article but none reported findings relevant to key question four) N=239 ears, 104 subjects	Time: not specified Place: A typical ear, nose, and throat clinic Affiliation: S1~ Dept. of Otolaryngology and Head and Neck Surgery, Hull Royal Infirmary, Hull, Great Britain Inclusion: • Age: 3 years– >=12 years (max not specified) • admitted for aspiration of the middle ear Exclusion: None Patient Characteristics: • 1% 3 years 6% 4 years 14% 5 years 23% 6 years 10% 7 years 13% 8 years 8% 9 years 10% 10 years 5% 11 years 10% ≥ 12 years	Comparisons: 1. Professional tympanometry Dx– : Compliance > .21 ml Dx+ : Compliance ≤ .21 ml Tympanocentesis (sedation unknown) GS– : Fluid absent GS+ : Fluid present 2. Professional tympanometry Dx– : Compliance > .25 ml Dx+ : Compliance ≤ .25 ml Tympanocentesis (sedation unknown) GS– : Fluid absent GS+ : Fluid present 3. Professional tympanometry Dx– : Compliance > .17 ml Dx+ : Compliance ≤ .17 ml Tympanocentesis (sedation unknown) GS– : Fluid absent GS+ : Fluid present 4. Professional tympanometry Dx– : Gradient > .04 ml/50 mm H2O Dx+ : Gradient ≤ .04 ml/ 50 mm H2O Tympanocentesis (sedation unknown) GS– : Fluid absent GS+ : Fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 75.96% (79/104) specificity 80.74% (109/135) PPV 75.24% (79/105) NPV 81.34% (109/134) accuracy 78.66% (188/239) prevalence 43.51% (104/239) Comparison 2: sensitivity 79.81% (83/104) specificity 74.81% (101/135) PPV 70.94% (83/117) NPV 82.79% (101/122) accuracy 76.99% (184/239) prevalence 43.51% (104/239) Comparison 3: sensitivity 69.23% (72/104) specificity 85.93% (116/135) PPV 79.12% (72/91) NPV 78.38% (116/148) accuracy 78.66% (188/239) prevalence 43.51% (104/239) Comparison 4: sensitivity 76.92% (80/104) specificity 85.19% (115/135) PPV 80.00% (80/100) NPV 82.73% (115/139) accuracy 81.59% (195/239) prevalence 43.51% (104/239) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1446 Haughton 1977			*Continued* 5. Professional tympanometry Dx– : Gradient > .03 ml/50 mm H2O Dx+ : Gradient ≤ .03 ml/ 50 mm H2O Tympanocentesis (sedation unknown) GS– : Fluid absent GS+ : Fluid present 6. Professional tympanometry Dx– : Gradient > .05 ml/50 mm H2O Dx+ : Gradient ≤ .05 ml/ 50 mm H2O Tympanocentesis (sedation unknown) GS– : Fluid absent GS+ : Fluid present 7. Professional tympanometry Dx– : Gradient > .1 ml/50 mm H2O Dx+ : Gradient ≤ .1 ml/ 50 mm H2O Tympanocentesis (sedation unknown) GS– : Fluid absent GS+ : Fluid present	*Continued* Unit of measure: Ear Comparison 5: sensitivity 73.08% (76/104) specificity 87.41% (118/135) PPV 39.18% (76/194) NPV 80.82% (118/146) accuracy 81.17% (194/239) prevalence 43.51% (104/239) Comparison 6: sensitivity 79.81% (83/104) specificity 81.48% (110/135) PPV 76.85% (83/108) NPV 83.97% (110/131) accuracy 80.75% (193/239) prevalence 43.51% (104/239) Comparison 7: sensitivity 87.50% (91/104) specificity 60.00% (81/135) PPV 62.76% (91/145) NPV 86.17% (81/94) accuracy 71.97% (172/239) prevalence 43.51% (104/239)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
3022 Johnson 1980	Cross-sectional Study Quality Score (0–6): 2 (110000) Examiner(s): • tympanometry (impedance bridge): audiologist • myringotomy: not specified *The primary aim of this study was to examine the changes in tympanogram types resulting from anesthesia. This is not a study focused on determining the sensitivity and/or specificity of tympanometry on detecting OME. Study Cohort: Children undergoing myringotomy with placement of ventillation tubes for recurrent AOM or persistent MEE N=61 subjects, 121 ears	Time: not specified Place: Ear, Nose, and Throat Clinic at the University of Utah Medical Center, Salt Lake City, UT Affiliation: Div. of Otolaryngology, Univ. of Utah Inclusion: • Age: 8 months–10 years • recurrent AOM or persistent MEE Exclusion: • perforated tympanic membrane (one ear excluded) Patient Characteristics: • mean age 3.5 years • male 40, female 21	Comparisons: 1. Professional tympanometry[a] Dx– : Normal tympanogram Dx+ : Flat tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 74.32% (55/74) specificity 75.61% (31/41) PPV 84.62% (55/65) NPV 62.00% (31/50) accuracy 74.78% (86/115) prevalence 64.35% (74/115)

[a] Findings excludes 4 ears that were found to have purulent effusion which is consistent with AOM and not OME and 2 where Professional tympanometry result 'unclassified.' According to Table 1 in the article those with Class 1 or 5 changes initially had normal (non-flat) tympanograms.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1600 Jonathan 1989	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • otoscopy (instrument not specified): not specified • tympanometry (instrument not specified): not specified • audiometry (instrument not specified): not specified • myringotomy: not specified *This study focused mainly on sonotubometry. Study Cohort: Children admitted for routine myringotomies (including in some cases adenoidectomy and/or tonsillectomy) *(a control group was studied however findings were not relevant to key question four) N=64 subjects, 126 ears	Time: not specified Place: Grp 1 were operative admissions, and Grp 2 were from a general pediatric ward. Affiliation: St. George's Hospital, Tooting, London Inclusion: • Age: 3–14 years • admitted for myringotomy Exclusion: • pre-existing perforations Patient Characteristics: • mean age 6.2 years, range 3–14 years; Grp 2 mean age 7 years, range 3–14 years • male 35, female 29;	Comparisons: 1. Non-pneumatic otoscopy[a] Dx– : normal appearance of the tympanic membrane Dx+ : abnormal appearance of the tympanic membrane Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry[a] Dx– : Tympanogram has any form other than a flat tracing Dx+ : Tympanogram is flat Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Audiometry – air conduction threshold [a] Dx– : ≤ 15 dB hearing threshold at all frequencies tested Dx+ : > 15 dB hearing threshold at all frequencies tested Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 100.00% specificity 28.00% PPV unknown NPV unknown accuracy 80.36% (90/112) prevalence 100.00% (112/112) Comparison 2: sensitivity 86.00% specificity 86.00% PPV unknown NPV unknown accuracy 85.71% (102/119) prevalence 100.00% (119/119) Comparison 3: sensitivity 90.00% specificity 52.00% PPV unknown NPV unknown accuracy 78.64% (81/103) prevalence 100.00% (103/103)

[a] Data neccesary to complete the results table were not provided.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1632 Karma 1989	Diagnostic study Study Quality Score (0–6): 3 (100101) Examiner(s): • pneumatic otoscopy: otolaryngologist in Tampere or pediatrician in Oulu • myringotomy (only peformed when MEE suspected by pneumatic otoscopy): not specified *OME definition:MEE present without other symptoms/signs of acute infection. Grp 1: Children seen by otolaryngologist in Tampere, Finland Grp 2: Children seen by pediatrician in Oulu, Finland N=2911 subjects, 11804 visits N1=1688 subjects, 5949 visits N2=1223 subjects, 5855 visits	Time: not specified Place: Outpatient Clinics of Otolaryngology at the Tampere or Pediatrics at the Oulu University Central Hospital, Finland Affiliation: Dept. of Clinical Sciences, Univ. of Tampere, Finland; National Public Health Institute, Helsinki, Finland Inclusion: • Age: 0.5–2.5 years • had any ear-related problem or suspicion of them and were brought to either the outpatient clinic in Tampere or Oulu Exclusion: • tympanostomy tubes in one or both ears at a visit • incomplete data Patient Characteristics: • Grp 1 age range 6–11 months 37%, 12–17 m 41%, 18–23 m 18%, 24–29 m 4%, 30–31 m 0.1%; Grp 2 age range 6–11 m 36%, 12–17 m 38%, 18–23 m 20%, 24–29 m 6%, 30–31 m 0.2%	Comparisons: 1. Grp 1 Signs/symptoms Dx– : TM color normal Dx+ : TM haemorrhagic Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Grp 1 Signs/symptoms Dx– : TM color normal Dx+ : TM strongly red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Grp 1 Signs/symptoms Dx– : TM color normal Dx+ : TM moderately red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear–related visits (1–2 years of follow-up) Comparison 1: sensitivity 0.25% (1/408) specificity 100.00% (185/185) PPV 100.00% (1/1) NPV 31.25% (185/592) accuracy 31.37% (186/593) prevalence 68.80% (408/593) Comparison 2: sensitivity 0.74% (3/408) specificity 99.46% (184/185) PPV 75.00% (3/4) NPV 31.24% (184/589) accuracy 31.53% (187/593) prevalence 68.80% (408/593) Comparison 3: sensitivity 2.21% (9/408) specificity 97.84% (181/185) PPV 69.23% (9/13) NPV 31.21% (181/580) accuracy 32.04% (190/593) prevalence 47.30% (408/593) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1632 Karma 1989			*Continued* 4. Grp 1 Signs/symptoms Dx– : TM color normal Dx+ : TM slightly red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 5. Grp 1[a] Non-pneumatic otoscopy Dx– : Tympanic membrane not red Dx+ : Tympanic membrane red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 6. Grp 1[a] Non-pneumatic otoscopy Dx– : Not red Dx+ : Distinctly red – hemorrhagic, strongly or moderately red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear–related visits Comparison 4: sensitivity 1.23% (5/408) specificity 97.84% (181/185) PPV 55.56% (5/9) NPV 30.99% (181/584) accuracy 31.37% (186/593) prevalence 68.80% (408/593) Comparison 5: sensitivity 4.66% (19/408) specificity 94.59% (175/185) PPV 65.52% (19/29) NPV 31.03% (175/564) accuracy 32.72% (194/593) prevalence 68.80% (408/593) LR+: 1.10 LR-: 1.00 Comparison 6: sensitivity 3.43% (14/408) specificity 97.30% (180/185) PPV 73.68% (14/19) NPV 31.36% (180/574) accuracy 32.72% (194/593) prevalence 68.80% (408/593) LR+: .50 LR-: 1.00 *Continued on next page*

[a] The raw numbers were calculated using a combination of information from tables 2, 3, and 5 in article.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1632 Karma 1989			*Continued* 7. Grp 1[a] Non-pneumatic otoscopy Dx– : TM not cloudy Dx+ : TM cloudy Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 8. Grp 1[a] Non-pneumatic otoscopy Dx– : TM color normal Dx+ : TM color abnormal Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 9. Grp 1[a] Non-pneumatic otoscopy Dx– : TM not bulging Dx+ : TM bulging Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear–related visits Comparison 7: sensitivity 92.89% (379/408) specificity 98.38% (182/185) PPV 99.21% (379/382) NPV 86.26% (182/211) accuracy 94.60% (561/593) prevalence 68.80% (408/593 LR+: 51.60 LR–: .07 Comparison 8: sensitivity 97.55% (398/408) specificity 92.97% (172/185) PPV 96.84% (398/411) NPV 94.51% (172/182) accuracy 96.12% (570/593) prevalence 68.80% (408/593) LR+: 13.90 LR–: .03 Comparison 9: sensitivity 45.10% (184/408) specificity 98.92% (183/185) PPV 98.92% (184/186) NPV 44.96% (183/407) accuracy 61.89% (367/593) prevalence 68.80% (408/593) LR+: 56.50 LR–: .60 *Continued on next page*

[a] The raw numbers were calculated using a combination of information from tables 2, 3, and 5 in article.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1632 Karma 1989			*Continued* 10. Grp 1[a] Non-pneumatic otoscopy Dx– : TM not retracted Dx+ : TM retracted Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 11. Grp 1[a] Non-pneumatic otoscopy Dx– : TM position normal Dx+ : TM position abnormal Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 12. Grp 1 Pneumatic otoscopy – unvalidated examiner Dx– : TM mobility normal Dx+ : TM mobility abnormal Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear–related visits Comparison 10: sensitivity 10.05% (41/408) specificity 95.14% (176/185) PPV 82.00% (41/50) NPV 32.41% (176/543) accuracy 36.59% (217/593) prevalence 68.80% (408/593) LR+: 2.00 LR–: .90 Comparison 11: sensitivity 55.39% (226/408) specificity 94.05% (174/185) PPV 95.36% (226/237) NPV 48.88% (174/356) accuracy 67.45% (400/593) prevalence 68.80% (408/593) LR+: 9.50 LR–: .50 Comparison 12: sensitivity 98.77% (403/408) specificity 90.27% (167/185) PPV 95.72% (403/421) NPV 97.09% (167/172) accuracy 96.12% (570/593) prevalence 68.80% (408/593) *Continued on next page*

[a] The raw numbers were calculated using a combination of information from tables 2, 3, and 5 in article.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1632 Karma 1989			*Continued* 13. Grp 1[a] Pneumatic otoscopy – examiner validation not specified Dx– : TM with normal mobility Dx+ : TM with impaired mobility Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 14. Grp 1 Pneumatic otoscopy – examiner validation not specified Dx– : TM with normal mobility Dx+ : TM mobility distinctly impaired Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 15. Grp 1 Pneumatic otoscopy – examiner validation not specified Dx– : TM mobility normal Dx+ : TM mobility slightly impaired Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear–related visits Comparison 13: sensitivity 98.77% (403/408) specificity 90.27% (167/185) PPV 95.72% (403/421) NPV 97.09% (167/172) accuracy 96.12% (570/593) prevalence 68.80% (408/593) LR+: 10.20 LR–: .01 Comparison 14: sensitivity 76.23% (311/408) specificity 97.30% (180/185) PPV 98.42% (311/316) NPV 64.98% (180/277) accuracy 82.80% (491/593) prevalence 68.80% (408/593) LR+: 30.50 LR–: .20 Comparison 15: sensitivity 22.30% (91/408) specificity 92.97% (172/185) PPV 87.50% (91/104) NPV 35.17% (172/489) accuracy 44.35% (263/593) prevalence 68.80% (408/593) LR+: 3.20 LR–: .80 *Continued on next page*

[a] The raw numbers were calculated using a combination of information from tables 2, 3, and 5 in article.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1632 Karma 1989			*Continued* 16. Grp 2 Signs/symptoms Dx– : TM color normal Dx+ : TM haemorrhagic Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 17. Grp 2 Signs/symptoms Dx– : TM color normal Dx+ : TM strongly red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 18. Grp 2 Signs/symptoms Dx– : TM color normal Dx+ : TM moderately red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear–related visits Comparison 16: sensitivity 5.80% (20/345) specificity 98.05% (151/154) PPV 86.96% (20/23) NPV 31.72% (151/476) accuracy 34.27% (171/499) prevalence 69.14% (345/499) Comparison 17: sensitivity 2.90% (10/345) specificity 99.35% (153/154) PPV 90.91% (10/11) NPV 31.35% (153/488) accuracy 32.67% (163/499) prevalence 29.06% (145/499) Comparison 18: sensitivity 2.90% (10/345) specificity 98.70% (152/154) PPV 83.33% (10/12) NPV 31.21% (152/487) accuracy 32.46% (162/499) prevalence 69.14% (345/499) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1632 Karma 1989			*Continued* 19. Grp 2 Signs/symptoms Dx– : TM color normal Dx+ : TM slightly red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 20. Grp 2 Non-pneumatic otoscopy Dx– : TM color normal Dx+ : Tympanic membrane red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 21. Grp 2[a] Non-pneumatic otoscopy Dx– : TM color normal Dx+ : TM distinctly red – hemorrhagic, strongly or moderately red Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear–related visits Comparison 19: sensitivity 8.99% (31/345) specificity 95.45% (147/154) PPV 81.58% (31/38) NPV 31.89% (147/461) accuracy 35.67% (178/499) prevalence 69.14% (345/499) Comparison 20: sensitivity 12.17% (42/345) specificity 92.21% (142/154) PPV 77.78% (42/54) NPV 31.91% (142/445) accuracy 36.87% (184/499) prevalence 69.14% (345/499) LR+: .20 LR–: 1.00 Comparison 21: sensitivity 11.30% (39/345) specificity 96.75% (149/154) PPV 88.64% (39/44) NPV 32.75% (149/455) accuracy 37.68% (188/499) prevalence 69.14% (345/499) LR+: 3.00 LR–: .90 *Continued on next page*

[a] The raw numbers were calculated using a combination of information from tables 2, 3, and 5 in article.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1632 Karma 1989			*Continued* 22. Grp 2[a] Non-pneumatic otoscopy Dx– : TM color normal Dx+ : TM cloudy Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 23. Grp 2[a] Non-pneumatic otoscopy Dx– : TM color normal Dx+ : TM color abnormal Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 24. Grp 2 [a] Non-pneumatic otoscopy Dx– : TM not bulging Dx+ : TM bulging Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear–related visits Comparison 22: sensitivity 68.99% (238/345) specificity 87.66% (135/154) PPV 92.61% (238/257) NPV 55.79% (135/242) accuracy 74.75% (373/499) prevalence 69.14% (345/499) LR+: 5.60 LR–: .40 Comparison 23: sensitivity 81.16% (280/345) specificity 79.22% (122/154) PPV 89.74% (280/312) NPV 65.24% (122/187) accuracy 80.56% (402/499) prevalence 69.14% (345/499) LR+: 4.00 LR–: .20 Comparison 24: sensitivity 18.26% (63/345) specificity 99.35% (153/154) PPV 98.44% (63/64) NPV 35.17% (153/435) accuracy 43.29% (216/499) prevalence 69.14% (345/499) LR+: 23.00 LR–: .80 *Continued on next page*

[a] The raw numbers were calculated using a combination of information from tables 2, 3, and 5 in article.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1632 Karma 1989			*Continued* 25. Grp 2[a] Non-pneumatic otoscopy Dx- : TM not retracted Dx+ : TM retracted Myringotomy (sedated) GS- : fluid absent GS+ : fluid present 26. Grp 2[a] Non-pneumatic otoscopy Dx– : TM position normal Dx+ : TM position abnormal Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 27. Grp 2 Pneumatic otoscopy – unvalidated examiner Dx– : TM mobility normal Dx+ : TM mobility abnormal Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear–related visits Comparison 25: sensitivity 32.17% (111/345) specificity 90.26% (139/154) PPV 88.10% (111/126) NPV 37.27% (139/373) accuracy 50.10% (250/499) prevalence 69.14% (345/499 LR+: 3.30 LR–: .80 Comparison 26: sensitivity 50.43% (174/345) specificity 90.20% (138/153) PPV 91.58% (174/190) NPV 44.66% (138/309) accuracy 62.53% (312/499) prevalence 69.14% (345/499) LR+: 4.90 LR–: .60 Comparison 27: sensitivity 93.62% (323/345) specificity 71.43% (110/154) PPV 88.01% (323/367) NPV 83.33% (110/132) accuracy 86.77% (433/499) prevalence 69.14% (345/499) *Continued on next page*

[a] The raw numbers were calculated using a combination of information from tables 2, 3, and 5 in article.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1632 Karma 1989			*Continued* 28. Grp 2[a] Pneumatic otoscopy – examiner validation not specified Dx– : TM mobility normal Dx+ : TM mobility impaired Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 29. Grp 2[a] Pneumatic otoscopy – examiner validation not specified Dx– : TM mobility normal Dx+ : TM mobility distinctly impaired Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 30. Grp 2[a] Pneumatic otoscopy – examiner validation not specified Dx– : TM mobility normal Dx+ : TM mobility slightly impaired Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear–related visits Comparison 28: sensitivity 93.62% (323/345) specificity 70.78% (109/154) PPV 87.77% (323/368) NPV 83.21% (109/131) accuracy 86.57% (432/499) prevalence 69.14% (345/499) LR+: 3.30 LR–: .09 Comparison 29: sensitivity 84.35% (291/345) specificity 85.71% (132/154) PPV 92.97% (291/313) NPV 70.97% (132/186) accuracy 84.77% (423/499) prevalence 69.14% (345/499) LR+: 5.90 LR–: .20 Comparison 30: sensitivity 9.28% (32/345) specificity 85.06% (131/154) PPV 58.18% (32/55) NPV 29.50% (131/444) accuracy 32.67% (163/499) prevalence 69.14% (345/499) LR+: .60 LR–: 1.10

[a] The raw numbers were calculated using a combination of information from tables 2, 3, and 5 in article.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1646 Kemaloglu 1999	Diagnostic study Study Quality Score (0–6): 3 (111000) Examiner(s): • otoscopy and pneumatic otoscopy: the first author • tympanometry (AZ7 electroacoustic impedancemeter [Interacoustics]): not specified • acoustic reflectometry (acoustic otoscope [ENT Medical Devices]) • paracentesis: not specified *Potential problem – only children with OME underwent paracentesis which was how the condition was diagnosed. The children without OME (controls) were diagnosed clinically, i.e., they did not undergo paracentesis. Grp 1: Ears with serous otitis media Grp 2: Normal Ears (age matched children who were free of any ENT problem during 3 month follow-up) N=156 subjects, 300 ears N1=81 subjects, 150 ears N2=75 subjects, 150 ears	Time: followed for 3 months (actual dates not specified) Place: Not specified Affiliation: Dept. of ENT, Faculty of Medicine, Gazi Univ., Ankara, Turkey Inclusion: • Grp 1~ears presented retracted TM in stage II or III with vascularization & dullness lasting for at least 3 months (effusion detected on paracentesis.) • Grp 2~free of any ENT problem during follow-up of 3 mos. Exclusion: • Grp 1~ears with tympanosclerosis or pseudomembrane of the TM and those with adhesive OM • Grp 1~subjects with recurrent otitis media who presented normal TM between intervals of AOM) Patient Characteristics: • Grp 1 mean age 6.17 years; Grp 2 mean age 6.42 years	Comparisons: 1. Grps 1 & 2 Professional tympanometry Dx– : A or C tracing on tympanogram Dx+ : B tracing on tympanogram Tympanocentesis (sedation unknown) GS– : fluid absent GS+ : fluid present 2. Grps 1 & 2 Acoustic reflectometry Dx– : Reflectivity < 5 Dx+ : Reflectivity ≥ 5 Tympanocentesis (sedation unknown) GS– : fluid absent GS+ : fluid present 3. Grps 1 & 2 Acoustic reflectometry Dx– : Curve angle > 75 degrees Dx+ : Curve angle ≤ 75 degrees Tympanocentesis (sedation unknown) GS– : fluid absent GS+ : fluid present 4. Grps 1 & 2 Acoustic reflectometry Dx– : Curve angle > 90 degrees Dx+ : Curve angle ≤ 90 degrees Tympanocentesis (sedation unknown) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 96.00% (144/150) specificity 92.00% (138/150) PPV 92.31% (144/156) NPV 95.83% (138/144) accuracy 94.00% (282/300) prevalence 50.00% (150/300) Comparison 2: sensitivity 65.33% (98/150) specificity 99.33% (149/150) PPV 98.99% (98/99) NPV 74.13% (149/201) accuracy 82.33% (247/300) prevalence 50.00% (150/300) Comparison 3: sensitivity 78.00% (117/150) specificity 99.33% (149/150) PPV 99.15% (117/118) NPV 81.87% (149/182) accuracy 88.67% (266/300) prevalence 50.00% (150/300) Comparison 4: sensitivity 97.33% (146/150) specificity 85.33% (128/150) PPV 86.90% (146/168) NPV 96.97% (128/132 accuracy 91.33% (274/300) prevalence 50.00% (150/300)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1650 Kennedy 1982	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • tympanometry (Madsen Z76): audiologist • myringotomy: not specified Study Cohort: Myringotomy or myringtomy with tubes performed on 75 ears suspected of having OME N=75 ears * number of subjects not specified	Time: not specified Place: Not specified Affiliation: Dept. of Otolaryngology, Geisinger Medical Clinic, Danville, PA Inclusion: • Ages: 9 months – 13 years • chronic or recurrent OME Exclusion: None Patient Characteristics: Not specified	Comparisons: 1. Professional tympanometry Dx– : Normal tympanogram Dx+ : Flat, rollover, or peaked negative pressure tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry Dx– : Normal, rollover, or peaked, or negative pressure tympanogram Dx+: Flat, only Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 100.00% (51/51) specificity 45.83% (11/24) PPV 79.69% (51/64) NPV 100.00% (11/11) accuracy 82.67% (62/75) prevalence 68.00% (51/75) Comparison 2: sensitivity 86.27% (44/51) specificity 87.50% (21/24) PPV 93.62% (44/47) NPV 75.00% (21/28) accuracy 86.67% (65/75) prevalence 68.00% (51/75)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1685 Koivunen 1997	Diagnostic study Study Quality Score (0–6): 3 (110001) Examiner(s): • tympanometry (MicroTymp [Welch Allyn]): a trained nurse • myringotomy: operating surgeon *Article does not state when the tympanogram was performed in relation to the myringotomy, thus it is not clear whether this study fits the criterion of the diagnostic test being performed ≤ 24hrs before myringotomy Study Cohort: Children referred for adenoidectomy, tympanostomy tube placement, or both procedures for recurrent otitis media or glue ear *(another group of children were studied but findings were not relevant to key question four) N=162 subjects, 314 ears	Time: Grp 1~8/1992–1/1993; Grp 2~two time periods during 1994 and 1995 (months not specified) Place: Oulu, Finland Affiliation: Depts. of Otolaryngology and Pediatrics, University of Oulu, Finland Inclusion: • referred for adenoidectomy, tympanostomy tube placement, or both procedures in the Dept. of Otolaryngology, in two time periods during 1994–95 Exclusion: • perforation or presence of tympanostomy tubes Patient Characteristics: • mean age 4.7 years, range 1 month-16 years; Grp 2 median age 34 months, range 7 months–8 years • male 130, female 76	Comparisons: 1. Portable tympanometer[a] Dx–: Type A and C tympanograms Dx+:Type B tympanograms Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Portable tympanometer[b] Dx– : Type A and C tympanograms Dx+:Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Portable tympanometer[c] Dx– : Type A and C tympanograms Dx+:Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 80.77% (42/52) specificity 92.97% (119/128) PPV 82.35% (42/51) NPV 92.25% (119/129) accuracy 89.44% (161/180) prevalence 28.89% (52/180) Comparison 2: sensitivity 71.43% (10/14) specificity 38.30% (18/47) PPV 25.64% (10/39) NPV 81.82% (18/22) accuracy 45.90% (28/61) prevalence 22.95% (14/61) Comparison 3: sensitivity 78.79% (52/66) specificity 78.29% (137/175) PPV 57.78% (52/90) NPV 90.73% (137/151) accuracy 78.42% (189/241) prevalence 27.39% (66/241) *Continued on next page*

[a] These results are for the "cooperative" group. These are combined results for children undergoing myringotomy for recurrent AOM and OME; results are not stratified on this variable.
[b] These results are for the "uncooperative" group. These are combined results for children undergoing myringotomy for recurrent AOM and OME; results are not stratified on this variable.
[c] These results are for the "cooperative" and "uncooperative" groups combined. Combined results for children undergoing myringotomy for recurrent AOM & OME; results aren't stratified on this variable.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1685 Koivunen 1997			*Continued* 4. Age: <24 months Portable tympanometer[a] Dx– : Type A and C tympanograms Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 5. Age: >=24 months Portable tympanometer[a] Dx– : Type A and C tympanograms Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear Comparison 4: sensitivity 78.00% specificity 91.00% PPV unknown NPV unknown accuracy unknown prevalence unknown Comparison 5: sensitivity 79.00% specificity 94.00% PPV unknown NPV unknown accuracy unknown prevalence unknown

[a] These are combined results for children undergoing myringotomy for recurrent AOM & OME; results are not stratified on this variable. Raw numbers were not provided for the results stratified by age.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1785 Lovette 1976	Diagnostic study Study Quality Score (0–6): 2 (100001) Examiner(s): • tympanometry (Model 1720 [Grason-Stadler]): nearly half by a third-year medical student • myringotomy: not specified Study Cohort: children involved in study to assess accuracy of detecting MEE by otoscopic tympanometric data alone * (another study was reported but findings were not relevant to key question four) N=21 subjects, 42 ears	Time: not specified Place: Affiliation: Tufts Univ., School of Engineering, Boston, MA; Dept. of Otolaryngology, Children's Hospital of Pittsburgh, PA Inclusion: • Age: Not specified • recurrent AOM or persistent MEE, or both • candidates for myringotomy and tympanostomy tube insertion Exclusion: None Patient Characteristics: Not Specified	Comparisons: 1. Portable tympanometer[a] Dx– : Admittance ≥ 0.25 mmhos Dx+ : Admittance < 0.25 mmhos Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 90.00% (27/30) specificity 58.33% (7/12) PPV 84.38% (27/32) NPV 70.00% (7/10) accuracy 80.95% (34/42) prevalence 71.43% (30/42)

[a] This study combines children undergoing myringotomy for recurrent AOM and OME. The results are not stratified on this variable.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1804 Macknin 1987	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • acoustic reflectometry: pediatrician • myringotomy: one of the staff otologists * This study was done on children with recurrent otitis media; they do not state whether this means recurrent OME or AOM. Study Cohort: Children who underwent myringotomy and tympanostomy tube placement for recurrent OM and who had reflectometry performed N=100 subjects, 198 ears	Time: not specified Place: Not specified Affiliation: Dept. of Pediatric and Adolescent Medicine and Dept. of Otolaryngology and Communicative Disorders, The Cleveland Clinic Foundation, Cleveland, OH Inclusion: • Age: 1–11 years (one 16 year old female) • underwent myringotomy and tympanostomy tube placement for recurrent OM and had reflectometry performed Exclusion: None Patient Characteristics: • mean age male 3.5 yrs and female 5.6 yrs, range 1–11 yrs except one 16 yr old • male 70, female 30	Comparisons: 1. Acoustic reflectometry[a] Dx– : Zero reflectivity units (RU) Dx+ : ≥ 1RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Acoustic reflectometry[a] Dx– : Zero reflectivity units (RU) Dx+ : ≥ 1 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Acoustic reflectometry[a] Dx– : < 2 reflectivity units (RU) Dx+ : ≥ 2 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 4. Acoustic reflectometry[a] Dx– : < 2 reflectivity units (RU) Dx+ : ≥ 2 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear (comparison specific) Comparison 1: Unit of measure: right ear sensitivity 98.51% (66/67) specificity 0.00% (0/33) PPV 66.67% (66/99) NPV 0.00% (0/1) accuracy 66.00% (66/100) prevalence 67.00% (67/100) Comparison 2: Unit of measure: left ear sensitivity 98.36% (60/61) specificity 2.70% (1/37) PPV 62.50% (60/96) NPV 50.00% (1/2) accuracy 62.24% (61/98) prevalence 62.24% (61/98) Comparison 3: Unit of measure: right ear sensitivity 97.01% (65/67) specificity 12.12% (4/33) PPV 69.15% (65/94) NPV 66.67% (4/6) accuracy 69.00% (69/100) prevalence 67.00% (67/100) Comparison 4: Unit of measure: left ear sensitivity 90.16% (55/61) specificity 5.41% (2/37) PPV 61.11% (55/90) NPV 25.00% (2/8) accuracy 58.16% (57/98) prevalence 62.24% (61/98) *Continued on next page*

[a] The values in the results table were calculated using the sensitivity and specificity provided in Table 2 in article; total "ns" of right and left ears examined (n=100 and n=98 respectively) and percentage of ears with fluid (67% and 62% respectively).

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1804 Macknin 1987			*Continued* 5. Acoustic reflectometry[a] Dx– : < 3 reflectivity units (RU) Dx+ : ≥ 3 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 6. Acoustic reflectometry[a] Dx– : < 3 reflectivity units (RU) Dx+ : ≥ 3 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 7. Acoustic reflectometry[a] Dx– : < 4 reflectivity units (RU) Dx+ : ≥ 4 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear (comparison specific) Comparison 5: Unit of measure: right ear sensitivity 89.55% (60/67) specificity 18.18% (6/33) PPV 68.97% (60/87) NPV 46.15% (6/13) accuracy 66.00% (66/100) prevalence 67.00% (67/100) Comparison 6: Unit of measure: left ear sensitivity 80.33% (49/61) specificity 21.62% (8/37) PPV 62.82% (49/78) NPV 40.00% (8/20) accuracy 58.16% (57/98) prevalence 62.24% (61/98) Comparison 7: Unit of measure: right ear sensitivity 82.09% (55/67) specificity 36.36% (12/33) PPV 72.37% (55/76) NPV 50.00% (12/24) accuracy 67.00% (67/100) prevalence 67.00% (67/100) *Continued on next page*

[a] The values in the results table were calculated using the sensitivity and specificity provided in Table 2 in article; total "ns" of right and left ears examined (n=100 and n=98 respectively) and percentage of ears with fluid (67% and 62% respectively).

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1804 Macknin 1987			*Continued* 8. Acoustic reflectometry[a] Dx– : < 4 reflectivity units (RU) Dx+ : ≥ 4 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 9. Acoustic reflectometry[a] Dx– : < 5 reflectivity units (RU) Dx+ : ≥ 5 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 10. Acoustic reflectometry[a] Dx– : < 5 reflectivity units (RU) Dx+ : ≥ 5 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 11. Acoustic reflectometry[a] Dx– : < 6 reflectivity units (RU) Dx+ : ≥ 6 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Comparison 8: Unit of measure: left ear sensitivity 70.49% (43/61) specificity 48.65% (18/37) PPV 69.35% (43/62) NPV 50.00% (18/36) accuracy 62.24% (61/98) prevalence 62.24% (61/98) Unit of measure: Ear (comparison specific) Comparison 9: right ear sensitivity 76.12% (51/67) specificity 60.61% (20/33) PPV 76.69% (51/64) NPV 55.56% (20/36) accuracy 71.00% (71/100) prevalence 67.00% (67/100) Comparison 10: left ear sensitivity 54.10% (33/61) specificity 62.16% (23/37) PPV 70.21% (33/47) NPV 45.10% (23/51) accuracy 57.14% (56/98) prevalence 62.24% (61/98) Comparison 11: right ear sensitivity 59.70% (40/67) specificity 72.73% (24/33) PPV 81.63% (40/49) NPV 47.06% (24/51) accuracy 64.00% (64/100) prevalence 67.00% (67/100) *Continued on next page*

[a] The values in the results table were calculated using the sensitivity and specificity provided in Table 2 in article; total "ns" of right and left ears examined (n=100 and n=98 respectively) and percentage of ears with fluid (67% and 62% respectively).

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1804 Macknin 1987			*Continued* 12. Acoustic reflectometry[a] Dx– : < 6 reflectivity units (RU) Dx+ : ≥ 6 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 13. Acoustic reflectometry[a] Dx– : < 7 reflectivity units (RU) Dx+ : ≥ 7 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 14. Acoustic reflectometry[a] Dx– : < 7 reflectivity units (RU) Dx+ : ≥ 7 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 15. Acoustic reflectometry[a] Dx– : < 8 reflectivity units (RU) Dx+ : ≥ 8 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear (comparison specific) Comparison 12: left ear sensitivity 36.07% (22/61) specificity 91.89% (34/37) PPV 88.00% (22/25) NPV 46.58% (34/73) accuracy 57.14% (56/98) prevalence 62.24% (61/98) Comparison 13: right ear sensitivity 47.76% (32/67) specificity 90.91% (30/33) PPV 91.43% (32/35) NPV 46.15% (30/65) accuracy 62.00% (62/100) prevalence 67.00% (67/100) Comparison 14: left ear sensitivity 26.23% (16/61) specificity 97.30% (36/37) PPV 94.12% (16/17) NPV 44.44% (36/81) accuracy 53.06% (52/98) prevalence 62.24% (61/98) Comparison 15: right ear sensitivity 17.91% (12/67) specificity 93.94% (31/33) PPV 85.71% (12/14) NPV 36.05% (31/86) accuracy 43.00% (43/100) prevalence 67.00% (67/100) *Continued on next page*

[a] The values in the results table were calculated using the sensitivity and specificity provided in Table 2 in article; total "ns" of right and left ears examined (n=100 and n=98 respectively) and percentage of ears with fluid (67% and 62% respectively).

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 1804 Macknin 1987			*Continued* 16. Acoustic reflectometry[a] Dx– : < 8 reflectivity units (RU) Dx+ : ≥ 8 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 17. Acoustic reflectometry[a] Dx– : < 9 reflectivity units (RU) Dx+ : ≥ 9 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 18. Acoustic reflectometry[a] Dx– : < 9 reflectivity units (RU) Dx+ : ≥ 9 RU Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear (comparison specific) Comparison 16: left ear sensitivity 3.28% (2/61) specificity 100.00% (37/37) PPV 100.00% (2/2) NPV 38.54% (37/96) accuracy 39.80% (39/98) prevalence 62.24% (61/98) Comparison 17: right ear sensitivity 1.50% (1/67) specificity 100.00% (33/33) PPV 100.00% (1/1) NPV 33.33% (33/99) accuracy 34.00% (34/100) prevalence 67.00% (67/100) Comparison 18: left ear sensitivity 0.00% (0/61) specificity 100.00% (37/37) PPV ------- (0/0) NPV 37.76% (37/98) accuracy 37.76% (37/98) prevalence 62.24% (61/98)

[a] The values in the results table were calculated using the sensitivity and specificity provided in Table 2 in article; total "ns" of right and left ears examined (n=100 and n=98 respectively) and percentage of ears with fluid (67% and 62% respectively).

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1817 Mains 1989	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • pneumatic otoscopy: a senior registrar and a senior house officer • myringotomy: not specified Study Cohort: Children admitted to Belfast City Hospital for myringotomy during a consecutive 5 month period N=114 subjects, 209 ears	Time: Patients entered study during a consecutive 5 month period (actual dates not specified) Place: Belfast City Hospital Affiliation: Not specified Inclusion: • admitted for myringotomy to the Belfast City Hospital during a 5-month period Exclusion: None Patient Characteristics: • age range 20 months–12 years • surgery indication primarily persistent middle ear effusion	Comparisons: 1. Examiner type: Observer[a] Pneumatic otoscopy – unvalidated examiner Dx– : TM mobile Dx+ : TM immobile Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Examiner type: Observer[b] Pneumatic otoscopy – unvalidated examiner Dx– : TM mobile Dx+ : TM immobile Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 87.93% (102/116) specificity 90.32% (84/93) PPV 91.89% (102/111) NPV 85.71% (84/98) accuracy 89.00% (186/209) prevalence 55.50% (116/209) Comparison 2: sensitivity 84.48% (98/116) specificity 87.10% (81/93) PPV 89.09% (98/110) NPV 81.82% (81/99) accuracy 85.65% (179/209) prevalence 55.50% (116/209)

[a] Observer described as 'Senior Registrar.'
[b] Observer described as 'Senior House Officer with just several months experience.'

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1837 Marchant 1986	Diagnostic study Study Quality Score (0–6): 2 (100001) Examiner(s): • pneumatic otoscope: one of two experienced otoscopists • tympanometry (otoadmittance meter, model 1723 version II [Grason-Stadler]): not specified • tympanocentesis: not specified *Tympanocentesis was only performed on 21 out of 73 infants. These are the only infants where a gold standard diagnostic procedure was performed. Select group where MD felt OM+ and child would benefit from knowing bug. Study Cohort : Infants 2–18 weeks old who underwent tympanocentesis * (Study Cohort was derived from an overall sample of 73 infants) N=21 subjects, 38 ears	Time: not specified Place: Pediatric Outpatient Clinic at Cleveland Metropolitan General Hospital Affiliation: Dept's of Pediatrics and Communication Services, Case Western Reserve University, OH; Division of Communication Disorders, Cleveland Metropolitan General Hospital, OH Inclusion: • Age: <5 months (all infants were 2–18 weeks of age) • attending well-child visit or otitis media was suspected after otoscopic examination by primary physician or nurse practitioner Exclusion: None Patient Characteristics: • age range 2–18 weeks	Comparisons: 1. Professional tympanometry Dx– : Peak susceptance > 0 mmho Dx+ : Peak susceptance ≤ 0 Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry Dx– : Ipsilateral acoustic reflex threshold ≤ 100 dB HL Dx+ : Ipsilateral acoustic reflex threshold absent or > 110 dB HL Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1[a]: sensitivity 94.29% (33/35) specificity[a] 100.00% (3/3) PPV 100.00% (33/33) NPV 60.00% (3/5) accuracy 94.74% (36/38) prevalence 92.11% (35/38) Comparison 2 [a]: sensitivity 97.14% (34/35) specificity 66.67% (2/3) PPV 97.14% (34/35) NPV 66.67% (2/3) accuracy 94.74% (36/38) prevalence 92.11% (35/38)

[a] The authors state that specificity could not be determined from the small number of ears without middle ear effusion on tympanocentesis.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
1936 Mitchell 1990	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): Unknown • tympanometry (Interacoustics AZ7 tympanometer [PCWerth]): • acoustic reflectometry (Acoustic otoscope [Endeco Medical]): • audiometry (Kamplex AC4 audiometer): • myringotomy: Study Cohort: Children with suspected glue ear were studied over a three-month period pure tone audiometry, tympanometry and acoustic reflectometry were attempted on each ear] N=50 subjects, 100 ears	Time: Three-month period (actual dates not specified) Place: Not specified Affiliation: Dept. of Otolaryngology, St. Thomas' Hospital, London Inclusion: • Age: 6 months–14 yrs • suspected glue ear Exclusion: None Patient Characteristics: • age range 6 months–14 years	Comparisons: 1. Professional tympanometry Dx– : Type A or C tympanogram Dx+ : Type B tympanogram Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present 2. Acoustic reflectometry Dx– : AR value ≤ 2 Dx+ : AR value > 2 Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present 3. Audiometry – air conduction threshold Dx– : Hearing threshold < 20dB Dx+ : Hearing threshold ≥ 20dB Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 87.69% (57/65) specificity 52.63% (10/19) PPV 86.36% (57/66) NPV 55.56% (10/18) accuracy 79.76% (67/84) prevalence 77.38% (65/84) Comparison 2: sensitivity 87.32% (62/71) specificity 59.26% (16/27) PPV 84.93% (62/73) NPV 64.00% (16/25) accuracy 79.59% (78/98) prevalence 72.45% (71/98) Comparison 3: sensitivity 80.39% (41/51) specificity 68.75% (11/16) PPV 89.13% (41/46) NPV 52.38% (11/21) accuracy 77.61% (52/67) prevalence 76.12% (51/67)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2012 Nozza 1992	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (GSI-33 Version I Middle Ear Analyzer [Grason-Stadler], 226 Hz probe tone): Grp 1 audiologist; Grp 2 nurse • myringotomy: surgeon • pneumatic otoscopy: nurse otoscopist validated at sensitivity and specificity >85% * This article includes 2 different studies: (1) examines tympanometry vs. myringotomy and the second examines validated pneumatic otoscopy vs. tympanometry; (2) represents the "general population." Grp 1: Ears in children undergoing myringotomy and tube surgery for treatment of chronic or recurrent OM Grp 2: Norms~Ears in children who were hospital outpatients and unscreened with respect to history of middle ear disease N=138 subjects, 171 ears N1=61 subjects, 111 ears N2=77 subjects, 144 ears	Time: not specified Place: Grp 1 Same-Day Surgery Unit of Children's Hospital of Pittsburgh, PA; Grp 2 children coming to outpatient allergy clinic Affiliation: Children's Hospital of Pittsburgh, PA; Dept.'s of Otolaryngology and Biostatistics, Univ. of Pittsburgh, PA Inclusion: • Age: Grp 1 1–8 years • Grp 1~undergoing myringotomy and tube surgery for treatment of chronic or recurrent OM • Grp 2~hospital outpatients who were unscreened with respect to history of middle ear disease (more representative of children in general population) Exclusion: None Patient Characteristics: • Grp 1 age range 1–8 years; Grp 2 mean age 9 years, range 3–16 years	Comparisons: 1. Grp 1[a] Professional tympanometry (AR alone) Dx– : acoustic reflex present Dx+ : acoustic reflex absent Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Grp 1[a] Professional tympanometry Dx– : Pressure gradient >0.1 Dx+ : Pressure gradient ≤0.1 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Grp 1[a] Professional tympanometry Dx– : Pressure gradient >0.2 Dx+ : Pressure gradient ≤0.2 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 4. Grp 1[a] Professional tympanometry Dx– : Peak compensated admittance >0.1 Dx+ : Peak compensated admittance ≤0.1 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	**Unit of measure: Ear** Comparison 1: sensitivity 88.16% (67/76) specificity 85.19% (23/27) PPV 94.37% (67/71) NPV 71.88% (23/32) accuracy 87.38% (90/103) prevalence 73.79% (76/103) Comparison 2: sensitivity 77.78% (63/81) specificity 90.00% (27/30) PPV 95.45% (63/66) NPV 60.00% (27/45) accuracy 81.08% (90/111) prevalence 72.97% (81/111) Comparison 3: sensitivity 91.36% (74/81) specificity 70.00% (21/30) PPV 89.16% (74/83) NPV 75.00% (21/28) accuracy 85.59% (95/111) prevalence 72.97% (81/111) Comparison 4: sensitivity 30.86% (25/81) specificity 96.67% (29/30) PPV 96.15% (25/26) NPV 34.12% (29/85) accuracy 48.65% (54/111) prevalence 72.97% (81/111) *Continued on next page*

[a] Values in the results table were determined by using the values provided in the text of this article which state no MEE group = 30 and MEE group = 81. Combining this information with the provided sensitivity and specificity allowed calculation of the numbers.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2012 Nozza 1992	*Continued*		*Continued* 5. Grp 1[a] Professional tympanometry Dx– : Peak compensated admittance > 0.2 Dx+ : Peak compensated admittance ≤ 0.2 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 6. Grp 1[a] Professional tympanometry Dx– : Peak compensated admittance > 0.3 Dx+ : Peak compensated admittance ≤ 0.3 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 7. Grp 1[a] Professional tympanometry Dx– : Peak compensated admittance > 0.4 Dx+ : Peak compensated admittance ≤ 0.4 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 5: sensitivity 55.56% (45/81) specificity 93.33% (28/30) PPV 95.74% (45/47) NPV 43.75% (28/64) accuracy 65.77% (73/111) prevalence 72.97% (81/111) Comparison 6: sensitivity 72.84% (59/81) specificity 80.00% (24/30) PPV 90.77% (59/65) NPV 52.17% (24/46) accuracy 74.77% (83/111) prevalence 72.97% (81/111) Comparison 7: sensitivity 81.48% (66/81) specificity 63.33% (19/30) PPV 85.71% (66/77) NPV 55.88% (19/34) accuracy 76.58% (85/111) prevalence 72.97% (81/111) *Continued on next page*

[a] Values in the above results table were determined by using the values provided in the text of this article which state no MEE group = 30 and MEE group = 81. Combining this information with the provided sensitivity and specificity allowed calculation of the numbers.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2012 Nozza 1992			*Continued* 8. Grp 2[a] Professional tympanometry Dx– : acoustic reflex present Dx+ : acoustic reflex absent Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 9. Grp 2[a] Professional tympanometry Dx– : Pressure gradient >0.1 Dx+ : Pressure gradient ≤ 0.1 Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 10. Grp 2[a] Professional tympanometry Dx– : Pressure gradient >0.2 Dx+ : Pressure gradient ≤ 0.2 Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 11. Grp 2[a] Professional tympanometry Dx– : Peak compensated admittance > 0.1 Dx+ : Peak compensated admittance ≤ 0.1 Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 8: sensitivity 87.50% (7/8) specificity 81.98% (91/111) PPV 25.93% (7/27) NPV 98.91% (91/92) accuracy 82.35% (98/119) prevalence 6.72% (8/119) Comparison 9: sensitivity 67.67% (6/9) specificity 100.00% (135/135) PPV 100.00% (6/6) NPV 97.83% (135/138) accuracy 97.92% (141/144) prevalence 6.25% (9/144) Comparison 10: sensitivity 77.78% (7/9) specificity 99.26% (134/135) PPV 87.50% (7/8) NPV 98.53% (134/136) accuracy 97.92% (141/144) prevalence 6.25% (9/144) Comparison 11: sensitivity 67.67% (6/9) specificity 100.00% (135/135) PPV 100.00% (6/6) NPV 97.83% (135/138) accuracy 97.92% (141/144) prevalence 6.25% (9/144) *Continued on next page*

[a] Values in the above results table were determined by using the values provided in the text of this article which state no MEE group = 135 and MEE group = 9. Combining this information with the provided sensitivity and specificity allowed calculation of the numbers.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2012 Nozza 1992			*Continued* 12. Grp 2[a] Professional tympanometry Dx– : Peak compensated admittance > 0.2 Dx+ : Peak compensated admittance ≤ 0.2 Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 13. Grp 2 [a] Professional tympanometry Dx– : Peak compensated admittance > 0.3 Dx+ : Peak compensated admittance ≤ 0.3 Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 14. Grp 2[a] Professional tympanometry Dx– : Peak compensated admittance > 0.4 Dx+ : Peak compensated admittance ≤ 0.4 Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present	*Continued* Unit of measure: Ear Comparison 12: sensitivity 77.78% (7/9) specificity 100.00% (135/135) PPV 100.00% (7/7) NPV 98.54% (135/137) accuracy 98.61% (142/144) prevalence 6.25% (9/144) Comparison 13: sensitivity 77.78% (7/9) specificity 97.78% (132/135) PPV 70.00% (7/10) NPV 98.51% (132/134) accuracy 96.53% (139/144) prevalence 6.25% (9/144) Comparison 14: sensitivity 77.78% (7/9) specificity 88.89% (120/135) PPV 31.82% (7/22) NPV 98.36% (120/122) accuracy 88.19% (127/144) prevalence 6.25% (9/144)

[a] Values in the above results table were calculated based on numbers provided under results on pg. 312 in article.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2013 Nozza 1994	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (GSI-33 Version I Middle Ear Analyzer [Grason-Stadler], 226 Hz probe tone): clinically certified and licensed audiologist • pneumatic otoscopy: Pediatric Nurse Practitioner validated at sensitivity 85% and specificity 71% • myringotomy: several surgeons This study combines results for children with chronic OM and recurrent otitis media; they do not state which of these categories represents OME or if either of them represents AOM. Grp 1: Myringotomy and Tympanostomy Group~Ears in children brought to operating room at Children's Hospital of Pittsburgh for M&T surgery Grp 2: Norms1~Ears studied by Nozza et al. N1=171 subjects, 249 ears N2=(subjects?), 130 ears	Time: not specified Place: Grp 1~Same-Day Surgery Unit of Children's Hospital of Pittsburgh, PA Affiliation: Children's Hospital of Pittsburgh, PA; Dept's of Otolaryngology and Biostatistics, Univ. of Pittsburgh, PA Inclusion: • Grp 1~Age: 1–12 years • Grp 1~chronic or recurrent OM • Grp 1~brought to operating room at Children's Hospital of Pittsburgh for M&T surgery Exclusion: None Patient Characteristics: • median age 3.8 years, range 1–12 years	Comparisons: 1. Grp 1 Diagnostic Method: unknown Dx– : tympanometric gradient > 0.2 Dx+ : tympanometric gradient ≤ 0.2 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Grp 1 Diagnostic Method: unknown Dx– : Tymp. Width ≤ 250 daPa Dx+ : Tymp. Width > 250 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 3. Grp 1[a] Pneumatic otoscopy – validated examiner Dx– : MEE absent Dx+ : MEE present Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 4. Grp 1 Professional tympanometry Dx– : Acoustic reflex present Dx+ : Acoustic reflex absent Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear (comp. 1–40) Unit of measure: Subject (comp. 41–43) Comparison 1: sensitivity 84.67% (116/137) specificity 61.61% (69/112) PPV 72.96% (116/159) NPV 76.67% (69/90) accuracy 74.30% (185/249) prevalence 55.02% (137/249) Comparison 2: sensitivity 84.67% (116/137) specificity 73.21% (82/112) PPV 79.45% (116/146) NPV 79.61% (82/103) accuracy 79.52% (198/249) prevalence 55.02% (137/249) Comparison 3: sensitivity 84.67% (116/137) specificity 71.43% (80/112) PPV 78.38% (116/148) NPV 79.21% (80/101) accuracy 78.71% (196/249) prevalence 55.02% (137/249) Comparison 4: sensitivity 85.48% (106/124) specificity 64.89% (61/94) PPV 76.26% (106/139) NPV 77.22% (61/79) accuracy 76.61% (167/218) prevalence 56.88% (124/218) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2013 Nozza 1994	*Continued*		*Continued* 5. Grp 1 Professional tympanometry Dx– : tympanometric gradient >0 Dx+ : tympanometric gradient ≤ 0 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 6. Grp 1 Professional tympanometry Dx– : tympanometric gradient >.1 Dx+ : tympanometric gradient ≤ 0.1 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 7. Grp 1 Professional tympanometry Dx– : tympanometric gradient >0.3 Dx+ : tympanometric gradient ≤0.3 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 8. Grp 1 Professional tympanometry Dx– : Peak admittance > 0 Dx+ : Peak admittance ≤ 0 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 5: sensitivity 23.36% (32/137) specificity 98.21% (110/112) PPV 94.12% (32/34) NPV 95.65% (110/115) accuracy 57.03% (142/249) prevalence 55.02% (137/249) Comparison 6: sensitivity 65.69% (90/137) specificity 91.07% (102/112) PPV 90.00% (90/100) NPV 68.46% (102/149) accuracy 77.11% (192/249) prevalence 55.02% (137/249) Comparison 7: sensitivity 92.70% (127/137) specificity 38.39% (43/112) PPV 64.80% (127/196) NPV 81.13% (43/53) accuracy 68.27% (170/249) prevalence 55.02% (137/249) Comparison 8: sensitivity 10.95% (15/137) specificity 98.21% (110/112) PPV 88.24% (15/17) NPV 47.41% (110/232) accuracy 50.20% (125/249) prevalence 55.02% (137/249) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2013 Nozza 1994			*Continued* 9. Grp 1 Professional tympanometry Dx– : Peak admittance >0.1 Dx+ : Peak admittance ≤0.1 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 10. Grp 1 Professional tympanometry Dx– : Peak admittance >0.2 Dx+ : Peak admittance ≤0.2 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 11. Grp 1 Professional tympanometry Dx– : Peak admittance >0.3 Dx+ : Peak admittance ≤0.3 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 12. Grp 1 Professional tympanometry Dx– : Peak admittance >0.4 Dx+ : Peak admittance ≤0.4 Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 9: sensitivity 27.01% (37/137) specificity 97.32% (109/112) PPV 92.50% (37/40) NPV 52.15% (109/209) accuracy 58.63% (146/249) prevalence 55.02% (137/249) Comparison 10: sensitivity 45.99% (63/137) specificity 91.96% (103/112) PPV 87.50% (63/72) NPV 58.19% (103/177) accuracy 66.67% (166/249) prevalence 55.02% (137/249) Comparison 11: sensitivity 70.07% (96/137) specificity 80.36% (90/112) PPV 81.36% (96/118) NPV 68.70% (90/131) accuracy 74.70% (186/249) prevalence 55.02% (137/249) Comparison 12: sensitivity 83.21% (114/137) specificity 68.75% (77/112) PPV 76.51% (114/149) NPV 77.00% (77/100) accuracy 76.71% (191/249) prevalence 55.02% (137/249) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2013 Nozza 1994			*Continued* 17. Grp 1 Professional tympanometry Dx– : tympanometric Width ≤300 daPa Dx+ : tympanometric Width >300daPa Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 18. Grp 1 Professional tympanometry Dx– : tympanometric Width ≤325 daPa Dx+ : tympanometric Width >325 daPa Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 19. Grp 1 Professional tympanometry Dx– : tympanometric Width ≤350 daPa Dx+ : tympanometric Width >350 daPa Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 20. Grp 1 Professional tympanometry Dx– : tympanometric Width ≤400 daPa Dx+ : tympanometric Width >400 daPa Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 17: sensitivity 76.64% (105/137) specificity 84.82% (95/112) PPV 86.07% (105/122) NPV 74.80% (95/127) accuracy 80.32% (200/249) prevalence 55.02% (137/249) Comparison 18: sensitivity 70.07% (96/137) specificity 88.39% (99/112) PPV 88.07% (96/109) NPV 70.71% (99/140) accuracy 78.31% (195/249) prevalence 55.02% (137/249) Comparison 19: sensitivity 61.31% (84/137) specificity 89.29% (100/112) PPV 87.50% (84/96) NPV 65.36% (100/153) accuracy 73.90% (184/249) prevalence 55.02% (137/249) Comparison 20: sensitivity 48.91% (67/137) specificity 96.43% (108/112) PPV 94.37% (67/71) NPV 60.67% (108/178) accuracy 70.28% (175/249) prevalence 55.02% (137/249) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2013 Nozza 1994			*Continued* 21. Grp 1 Professional tympanometry Dx– : Acoustic reflex present Dx+ : Acoustic reflex absent Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 22. Grp 1 Professional tympanometry Dx– : Tympanometric gradient >0 Dx+ : Tympanometric gradient ≤0 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 23. Grp 1 Professional tympanometry Dx– : Tympanometric gradient >0.1 Dx+ : Tympanometric gradient ≤0.1 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 24. Grp 1 Professional tympanometry Dx– : Tympanometric gradient >0.2 Dx+ : Tympanometric gradient ≤0.2 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 21: sensitivity 76.61% (95/124) specificity 76.60% (72/94) PPV 81.20% (95/117) NPV 71.29% (72/101) accuracy 76.61% (167/218) prevalence 56.88% (124/218) Comparison 22: sensitivity 23.36% (32/137) specificity 100.00% (112/112) PPV 100.00% (32/32) NPV 95.73% (112/117) accuracy 57.83% (144/249) prevalence 55.02% (137/249) Comparison 23: sensitivity 64.23% (88/137) specificity 93.75% (105/112) PPV 92.63% (88/95) NPV 68.18% (105/154) accuracy 77.51% (193/249) prevalence 55.02% (137/249) Comparison 24: sensitivity 84.67% (116/137) specificity 65.18% (73/112) PPV 74.84% (116/155) NPV 77.66% (73/94) accuracy 75.90% (189/249)) prevalence 55.02% (137/249) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2013 Nozza 1994			*Continued* 25. Grp 1 Professional tympanometry Dx– : Tympanometric gradient >0.3 Dx+ : Tympanometric gradient ≤0.3 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 26. Grp 1 Professional tympanometry Dx– : Peak admittance >0 Dx+ : Peak admittance ≤0 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 27. Grp 1 Professional tympanometry Dx– : Peak admittance > 0.1 Dx+ : Peak admittance ≤0.1 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 28. Grp 1 Professional tympanometry Dx– : Peak admittance >0.2 Dx+ : Peak admittance ≤0.2 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 25: sensitivity 92.70% (127/137) specificity 32.14% (36/112) PPV 62.56% (127/203) NPV 78.26% (36/46) accuracy 65.46% (163/249) prevalence 55.02% (137/249) Comparison 26: sensitivity 10.95% (15/137) specificity 100.00% (112/112) PPV 100.00% (15/15) NPV 47.86% (112/234) accuracy 51.00% (127/249) prevalence 55.02% (137/249) Comparison 27: sensitivity 26.28% (36/137) specificity 100.00% (112/112) PPV 100.00% (36/36) NPV 99.12% (112/113) accuracy 59.44% (148/249) prevalence 55.02% (137/249) Comparison 28: sensitivity 46.72% (64/137) specificity 97.32% (109/112) PPV 95.52% (64/67) NPV 59.89% (109/182) accuracy 69.48% (173/249) prevalence 55.02% (137/249) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2013 Nozza 1994			*Continued* 29. Grp 1 Professional tympanometry Dx– : Peak admittance >0.3 Dx+ : Peak admittance ≤0.3 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 30. Grp 1 Professional tympanometry Dx– : Peak admittance >0.4 Dx+ : Peak admittance ≤0.4 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 31. Grp 1 Professional tympanometry Dx– : Tympanometric Width ≤150 daPa Dx+ : Tympanometric Width >150 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 32. Grp 1 Professional tympanometry Dx– : Tympanometric Width ≤200 daPa Dx+ : Tympanometric Width >200 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 29: sensitivity 70.07% (96/137) specificity 84.82% (95/112) PPV 84.96% (96/113) NPV 69.85% (95/136) accuracy 76.71% (191/249) prevalence 55.02% (137/249) Comparison 30: sensitivity 76.74% (105/137) specificity 66.07% (74/112) PPV 73.43% (105/143) NPV 69.81% (74/106) accuracy 71.89% (179/249) prevalence 55.02% (137/249) Comparison 31: sensitivity 94.89% (130/137) specificity 28.57% (32/112) PPV 61.90% (130/210) NPV 82.05% (32/39) accuracy 65.06% (162/249) prevalence 55.02% (137/249) Comparison 32: sensitivity 89.05% (122/137) specificity 50.89% (57/112) PPV 68.93% (122/177) NPV 79.17% (57/72) accuracy 71.89% (179/249) prevalence 55.02% (137/249) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2013 Nozza 1994			*Continued* 33. Grp 1 Professional tympanometry Dx– : Tymp. Width ≤225 daPa Dx+ : Tymp. Width >225 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 34. Grp 1 Professional tympanometry Dx– : Tymp. Width ≤275 daPa Dx+ : Tymp. Width >275 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 35. Grp 1 Professional tympanometry Dx– : Tymp. Width ≤300 daPa Dx+ : Tymp Width >300 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 36. Grp 1 Professional tympanometry Dx– : Tymp. Width ≤325 daPa Dx+ : Tymp. Width >325 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 33: sensitivity 86.13% (118/137) specificity 61.61% (69/112) PPV 73.29% (118/161) NPV 78.41% (69/88) accuracy 75.10% (187/249) prevalence 55.02% (137/249) Comparison 34: sensitivity 78.83% (108/137) specificity 85.71% (96/112) PPV 87.10% (108/124) NPV 76.80% (96/125) accuracy 81.93% (204/249) prevalence 55.02% (137/249) Comparison 35: sensitivity 78.83% (103/137) specificity 86.61% (97/112) PPV 87.29% (103/118) NPV 74.05% (97/131) accuracy 80.32% (200/249) prevalence 55.02% (137/249) Comparison 36: sensitivity 70.07% (96/137) specificity 93.75% (105/112) PPV 93.20% (96/103) NPV 71.92% (105/146) accuracy 80.72% (201/249) prevalence 55.02% (137/249) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2013 Nozza 1994			*Continued* 37. Grp 1 Professional tympanometry Dx– : Tymp. Width ≤350 daPa Dx+ : Tymp. Width >350 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 38. Grp 2 Professional tympanometry Dx– : Tymp. Width ≤150 daPa Dx+ : Tymp. Width >150 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 39. Grp 2 Professional tympanometry Dx– : Tymp. Width ≤200 daPa Dx+ : Tymp. Width >200 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 40. Grp 2 Professional tympanometry Dx– : Tymp. Width ≤250 daPa Dx+ : Tymp. Width >250 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 37: sensitivity 62.04% (85/137) specificity 94.64% (106/112) PPV 93.41% (85/91) NPV 67.09% (106/158) accuracy 76.71% (191/249) prevalence 55.02% (137/249) Comparison 38: sensitivity 88.89% (8/9) specificity 91.97% (126/137) PPV 47.06% (8/17) NPV 99.21% (126/127) accuracy 93.06% (134/144) prevalence 6.25% (9/144) Comparison 39: sensitivity 77.78% (7/9) specificity 99.26% (134/135) PPV 87.50% (7/8) NPV 98.53% (134/136) accuracy 97.92% (141/144) prevalence 6.25% (9/144) Comparison 40: sensitivity 77.78% (7/9) specificity 100.00% (135/135) PPV 100.00% (7/7) NPV 98.54% (135/137) accuracy 98.61% (142/144) prevalence 6.25% (9/144) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2013 Nozza 1994			*Continued* 41. Grp 2 Professional tympanometry Dx– : Tymp. Width ≤150 daPa Dx+ : Tymp. Width >150 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 42. Grp 2 Professional tympanometry Dx– : Tymp. Width ≤ 200 daPa Dx+ : Tymp. Width >200 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 43. Grp 2 Professional tympanometry Dx– : Tymp. Width ≤250 daPa Dx+ : Tymp. Width >250 daPa Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present	*Continued* Unit of measure: Subjects Comparison 41: sensitivity 88.89% (8/9) specificity 79.41% (54/68) PPV 36.36% (8/22) NPV 98.18% (54/55) accuracy 80.52% (62/77) prevalence 11.69% (9/77) Comparison 42: sensitivity 88.89% (8/9) specificity 89.71% (61/68) PPV 53.33% (8/15) NPV 98.39% (61/62) accuracy 89.61% (69/77) prevalence 11.69% (9/77) Comparison 43: sensitivity 88.89% (8/9) specificity 92.65% (63/68) PPV 61.54% (8/13) NPV 98.44% (63/64) accuracy 92.21% (71/77) prevalence 11.69% (9/77)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2048 Orchik 1978	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (electroacoustic impedance bridge, Madsen 70–72): performer not specified • myringotomy: surgeon Study Cohort: Ears of patients who underwent myringotomy for suspected serous otitis media N=76 ears * number of subjects not specified	Time: not specified Place: Not specified Affiliation: Dept. of Audiology and Speech Pathology, Memphis State Univ., TN; East Texas Rehabilitation Center Inclusion: • Age: 6 months–14 years • underwent myringotomy for suspected serous otitis media Exclusion: None Patient Characteristics: • mean age 4.5 years, range 6 months–14 years	Comparisons: 1. Portable tympanometer[a] Dx– : Type A and C tympanograms Dx+ : Type B tympanogram Myringotomy (sedation unknown) GS– : fluid absent or minimal GS+ : fluid present in moderate or large amounts 2. Professional tympanometry[b] Dx– : Acoustic reflex present Dx+ : Acoustic reflex absent Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present 3. not included 4. Professional tympanometry Dx– : Type A & C1&C3 tympanogram Dx+ : Types B & C2 tympanograms Myringotomy (sedation unknown) GS– : fluid absent or minimal GS+ : fluid present in moderate or large amounts *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 53.85% (21/39) specificity 94.59% (35/37) PPV 91.30% (21/23) NPV 66.04% (35/53) accuracy 73.68% (56/76) prevalence 51.32% (39/76) Comparison 2: sensitivity 89.74% (35/39) specificity 78.38% (29/37) PPV 81.40% (35/43) NPV 87.88% (29/33) accuracy 84.21% (64/76) prevalence 51.32% (39/76) Comparison 3: (not included-same as C1) Comparison 4: sensitivity 79.49% (31/39) specificity 86.49% (32/37) PPV 86.11% (31/36) NPV 80.00% (32/40) accuracy 82.89% (63/76) prevalence 51.32% (39/76) *Continued on next page*

[a] In this results table the type "c" tympanograms are included as normals.
[b] Values derived from false positive % and false negative % reported in Table VI for acoustic reflex and total numbers of normal and abnormal ears as reported in Table 2.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2048 Orchik 1978			*Continued* 5. Professional tympanometry[a] Dx– : acoustic reflex present at 500, 2,000, or 4,000 Hz Dx+ : acoustic reflex absent at 500, 2000, or 4000 Hz Myringotomy (sedation unknown) GS– : fluid absent or minimal GS+ : fluid moderate/large amounts 6. Professional tympanometry [a] Dx– : acoustic reflex present at 1000Hz Dx+ : acoustic reflex absent at 1000 Hz Myringotomy (sedation unknown) GS– : fluid absent or minimal GS+ : fluid moderate/large amounts 7. Quantitative tympanometry Dx– : Static compliance ≥.28cc Dx+ : Static compliance <.28 Myringotomy (sedation unknown) GS– : fluid absent or minimal GS+ : fluid present in moderate or large amounts	*Continued* Unit of measure: Ear Comparison 5: sensitivity 87.18% (34/39) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence 51.32% (39/76) Comparison 6: sensitivity 89.74% (35/39) specificity unknown PPV unknown NPV unknown accuracy unknown prevalence 51.32% (39/76) Comparison 7: sensitivity 89.74% (35/39) specificity 40.54% (15/37) PPV 61.40% (35/57) NPV 78.95% (15/19) accuracy 65.79% (50/76) prevalence 51.32% (39/76)

[a] Specificity data might be abstractable from this article for these cutpoints.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2049 Orchik 1978	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (electroacoustic impedance bridge): not specified • myringotomy: surgeon Study Cohort: Patients who were undergoing myringotomy for suspected serous otitis media N=75 subjects, 142 ears	Time: not specified Place: not specified Affiliation: Div. of Communication Disorder, North Texas State Univ.; J.W. Dunn, MD, and Associates, Denton TX Inclusion: • Age: Not specified • undergoing myringotomy for suspected serous otitis media Exclusion: None Patient Characteristics: Not Specified	Comparisons: 1. Professional tympanometry[a] Dx– : Type A and As tympanograms Dx+ : Type B tympanograms Myringotomy (sedation unknown) GS– : fluid absent or minimal GS+ : fluid present in moderate or large amounts 2. Professional tympanometry[b] Dx– : Type A, As, C, and Cs tympanograms Dx+ : Type B tympanograms Myringotomy (sedation unknown) GS– : fluid absent or minimal GS+ : fluid present in moderate or large amounts 3. Professional tympanometry[c] Dx– : Types A and As Dx+ : Types B, C, and Cs Myringotomy (sedation unknown) GS– : fluid absent or minimal GS+ : fluid moderate/large amounts	Unit of measure: Ear Comparison 1: sensitivity 75.41% (46/61) specificity 81.25% (26/32) PPV 88.46% (46/52) NPV 63.41% (26/41) accuracy 77.42% (72/93) prevalence 65.59% (61/93) Comparison 2: sensitivity 54.76% (46/84) specificity 89.66% (52/58) PPV 88.46% (46/52) NPV 57.78% (52/90) accuracy 69.01% (98/142) prevalence 59.15% (84/142) Comparison 3: sensitivity 82.14% (69/84) specificity 44.83% (26/58) PPV 68.32% (69/101) NPV 63.41% (26/41) accuracy 66.90% (95/142) prevalence 59.15% (84/142)

[a] Excluded data presented for type "c" tympanograms. Classified ears with "none" and "minimal" effusion as = no effusion. Classified ears with "moderate" or "impacted" as = +effusion.
[b] Type "c" and "cs" tympanograms are included as "normals." Classified ears with "none" and "minimal" effusion as = no effusion. Classified ears with "moderate" or "impacted" as = +effusion.
[c] Type "c" and "cs" tympanograms are included as "abnormals." Classified ears with "none" and "minimal" effusion as = no effusion. Classified ears with "moderate" or "impacted" as = +effusion.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2050 Orchik 1980	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (electroacoustic impedance bridge, Madsen ZO-72): not specified • myringotomy: surgeon Study Cohort: Ears of patients undergoing myringotomy for suspected serous otitis media N=76 ears * number of subjects not specified	Time: not specified Place: not specified Affiliation: Memphis State Univ, TN; J.W. Dunn, MD, and Associates, TX Inclusion: • Age: 6 months–14 years • undergoing myringotomy for suspected serous otitis media, usually for recurrent OM or failure of single OM episode to respond to medical treatment Exclusion: None Patient Characteristics: • mean age 4.5 years, range 6 months–14 years	Comparisons: 1. Professional tympanometry[a] Dx– : Type A tympanogram Dx+ : Type B tympanpgram Myringotomy (sedation unknown) GS– : Fluid absent or minimal GS+ : Fluid present in moderate or large amounts 2. Professional tympanometry[b] Dx– : Type A and C tympanograms Dx+ : Type B tympanograms Myringotomy (sedation unknown) GS– : Fluid absent or minimal GS+ : Fluid present in moderate or large amounts 3. Professional tympanometry[b] Dx– : Type A tympanograms Dx+ : Type B and C tympanograms Myringotomy (sedation unknown) GS– : Fluid absent or minimal GS+ : Fluid present in moderate or large amounts *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 87.50% (21/24) specificity 92.31% (24/26) PPV 91.30% (21/23) NPV 88.89% (24/27) accuracy 90.00% (45/50) prevalence 48.00% (24/50) Comparison 2: sensitivity 53.85% (21/39) specificity 94.59% (35/37) PPV 91.30% (21/23) NPV 66.04% (35/53) accuracy 73.68% (56/76) prevalence 51.32% (39/76) Comparison 3: sensitivity 92.31% (36/39) specificity 64.86% (24/37) PPV 73.47% (36/49) NPV 88.89% (24/27) accuracy 78.95% (60/76) prevalence 51.32% (39/76) *Continued on next page*

[a] Excluded data presented for type "c" tympanograms. Ears with "none" and "minimal" effusion classified as = no effusion; ears with "moderate" or "impacted" as = +effusion.
[b] Type "c" tympanograms are included as "normals." Ears with "none" and "minimal" effusion classified as = no effusion; ears with "moderate" or "impacted" as = +effusion.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2050 Orchik 1980			*Continued* 4. Professional tympanometry Dx– : Acoustic refl. present @500 Hz Dx+ : Acoustic refl. absent @500 Hz Myringotomy (sedation unknown) GS– : Fluid absent or minimal GS+ : Fluid present in moderate or large amounts 5. Professional tympanometry Dx– : Acoustic reflex present at 1000 Hz Dx+ : Acoustic reflex absent at 1000 Hz Myringotomy (sedation unknown) GS– : Fluid absent or minimal GS+ : Fluid present in moderate or large amounts 6. Professional tympanometry Dx– : Acoustic reflex present at 2000 Hz Dx+ : Acoustic reflex absent at 2000 Hz Myringotomy (sedation unknown) GS– : Fluid absent or minimal GS+ : Fluid present in moderate or large amounts *Continued on next page*	*Continued* Unit of measure: Ear Comparison 4: sensitivity 87.18% (34/39) specificity 70.27% (26/37) PPV 75.56% (34/45) NPV 83.87% (26/31) accuracy 78.95% (60/76) prevalence 51.32% (39/76) Comparison 5: sensitivity 89.74% (35/39) specificity 72.97% (27/37) PPV 77.78% (35/45) NPV 87.10% (27/31) accuracy 81.58% (62/76) prevalence 51.32% (39/76) Comparison 6: sensitivity 87.18% (34/39) specificity 75.68% (28/37) PPV 79.07% (34/43) NPV 84.85% (28/33) accuracy 81.58% (62/76) prevalence 51.32% (39/76) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2050 Orchik 1980			*Continued* 7. Professional tympanometry Dx– : Acoustic reflex present at 4000 Hz Dx+ : Acoustic reflex absent at 4000 Hz Myringotomy (sedation unknown) GS– : Fluid absent or minimal GS+ : Fluid present in moderate or large amounts 8. Quantitative tympanometry Dx– : Static compliance ≥ .28 cc Dx+ : Static compliance < .28 cc Myringotomy (sedation unknown) GS– : Fluid absent or minimal GS+ : Fluid present in moderate or large amounts	*Continued* Unit of measure: Ear Comparison 7: sensitivity 87.18% (34/39) specificity 75.68% (28/37) PPV 79.07% (34/43) NPV 84.85% (28/33) accuracy 81.58% (62/76) prevalence 51.32% (39/76) Comparison 8: sensitivity 89.74% (35/39) specificity 40.54% (15/37) PPV 61.40% (35/57) NPV 78.95% (15/19) accuracy 65.79% (50/76) prevalence 51.32% (39/76)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2055 Ovesen 1993	N/A Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (Tympan-O-scope model zs 330 [Madsen Electronics]): ENT physician • otomicroscopy: not specified • myringotomy: two other ENT specialists Study Cohort: Children with unilateral or bilateral SOM N=222 subjects, 440 ears	Time: not specified Place: Not specified Affiliation: Ear Nose and Throat Dept., Aarhus Univ. Hospital, Aarhus C, Denmark Inclusion: • Age: 0.8–14.8 years • two of the following three criteria: otomicroscopical findings consistent with SOM during 3 months, >20 dB hearing impairment, and/or adenoid symptoms Exclusion: None Patient Characteristics: • mean age 4.1 years, 0.8–14.8 years • male 132, female 88	Comparisons: 1. Professional tympanometry[a] Dx– : Type A tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry[b] Dx– : Type A/C1/C2 tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Professional tympanometry[c] Dx– : Type A or C1 tympanogram Dx+ : Type B or C2 tympanogram Myringotomy (sedated) GS– : Fluid absent GS+ : Fluid present 4. Professional tympanometry[d] Dx– : Type A tympanogram Dx+ : Types B/C1/C2 tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 99.36% (310/312) specificity 48.15% (13/27) PPV 95.68% (310/324) NPV 86.67% (13/15) accuracy 95.28% (323/339) prevalence 92.04% (312/339) Comparison 2: sensitivity 90.64% (310/342) specificity 72.55% (37/51) PPV 95.68% (310/324) NPV 53.62% (37/69) accuracy 88.30% (347/393) prevalence 87.02% (342/393) Comparison 3: sensitivity 94.44% (323/342) specificity 52.94% (27/51) PPV 93.08% (323/347) NPV 58.70% (27/46) accuracy 89.06% (350/393) prevalence 87.02% (342/393) Comparison 4: sensitivity 99.42% (340/342) specificity 25.49% (13/51) PPV 89.95% (340/378) NPV 86.67% (13/15) accuracy 89.82% (353/393) prevalence 87.02% (342/393)

[a] Excludes cases where no operation was performed. It also excludes cases where an operation was performed and the tympanogram was type C1 or C2.
[b] Excludes cases where no operation was performed. Type C1 and C2 are classified as “normal” in the above results table.
[c] Excludes cases where no operation was performed. Type A and C1 are classified as “normal”and type B and C2 are classified as “abnormal” in the above results table.
[d] Excludes cases where no operation was performed. Type A is classified as “normal”and types B, C1, and C2 are classified as “abnormal” in the above results table.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2058 Oyiborhoro 1987	Diagnostic study Study Quality Score (0–6): 3 (100101) Examiner(s): • acoustic reflectometry (acoustic otoscope [Endeco Medical]): not specified • pneumatic otoscopy: otolaryngologist validated with sensitivity 97.73% and specificity 90.90% • tympanometry diagnostic impedance audiometer Madsen ZS 77 MB): not specified • audiometry (Beltone clinical audiotry model 200-C): not specified • myringotomy Grp 1: Children who did not have MEE Grp 2: Children who had MEE Grp 3: Subgroup of children in group2 who underwent myringotomy (44 pathologic ears) N1=100 subjects, 200 ears N2=100 subjects, 175 ears N3=23 subjects, 44 ears	Time: Referred over a 10-month period (actual dates not specified) Place: Subjects drawn from the ENT clinic at the Dr. Martin Luther King, Jr. Health Center, Bronx, NY Affiliation: Columbia Univ., NY; VA Medical Center, Cleveland, OH; as above Inclusion: • Age: 3–12 years • Grp 1: normal hearing and middle ear function based upon results from otoscopy, tympanometry and pure tone audiometry) • Grp 2: pure tone thresholds ≥20dB HL (250–8000 Hz); air-bone gaps of at least 15 dB HL (250–4000 Hz); type B tympanograms in the ears studied w/no sensorineural components; or demonstrated hearing and immittance abnormalities at least two times in ears studied during the 10-month period of investigation. Exclusion: • Grp 1~history suggested the existence of any known middle ear pathology Patient Characteristics: • mean age 8.7 years, range 3–12 years • male 90, female 110	Comparisons: 1. Grps 1 & 2 combined Acoustic reflectometry Dx– : Acoustic otoscope reflectivity<4 Dx+ : Acoustic otoscope reflect. ≥4 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 2. Grps 1 & 2 combined Acoustic reflectometry Dx– : Acoustic otoscope reflectivity<3 Dx+ : Acoustic otoscope reflect. ≥3 Validated pneumatic otoscopy GS– : MEE absent GS+ : MEE present 3. Grp 3[a] Acoustic reflectometry Dx– : Acoustic otoscope reflectivity<4 Dx+ : Acoustic otoscope reflect. ≥4 Myringotomy (sedation unknown) GS– : Fluid absent GS+ : Fluid present	Unit of measure: Ear Comparison 1: sensitivity 93.14% (163/175) specificity 83.00% (166/200) PPV 82.74% (163/197) NPV 93.26% (166/178) accuracy 87.73% (329/375) prevalence 46.67% (175/375) Comparison 2: sensitivity 97.71% (171/175) specificity 65.50% (131/200) PPV 71.25% (171/240) NPV 97.04% (131/135) accuracy 80.53% (302/375) prevalence 46.67% (175/375) Comparison 3: sensitivity 95.45% (42/44) specificity - (0/0) PPV 100.00% (42/42) NPV - (0/2) accuracy 95.45% (42/44) prevalence 100.00% (44/44)

[a] All children in this group had fluid on myringotomy, thus specificity cannot be determined for this group.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2118 Paradise 1976	Diagnostic study This ID number includes two separate studies Study Quality Score (0–6): STUDY 1: 4 (111001) STUDY 2: 5 (111101) Examiner(s): • tympanometry (Madsen Electroacoustic Impedance Meter, Model ZO 70 [Madsen Electronics]): audiologist • pneumatic otoscopy: pediatrician • myringotomy: by or under supervision of one of two otolaryngologists Study Cohort: infants and children who had been scheduled for myringotomy and insertion of tympanostomy tubes because of recurrent AOM or persistent MEE, or both * (group of infants and children not receiving myringotomy were studied but findings were not relevant to key question four) N=107 subjects, 214 ears	Time: 5/25/1972–5/9/1974 Place: Grp 1 not specified: Grp 2 various outpatient waiting areas at Children's Hospital of Pittsburgh, PA Affiliation: Dept. of Pediatrics, Community Med., and Otolaryngology, and Speech and Theatre Arts Cleft Palate Center, Univ. of Pittsburgh; Ambulatory Care Center, Children's Hospital of Pittsburgh Inclusion: • Age: 10 days–5 years Grp 1: • scheduled for myringotomy and insertion of tubes due to recurrent AOM or persistent MEE or both Grp 2: • not scheduled for myringotomy • in various outpatient waiting areas of the Children's Hospital of Pittsburgh; available and parents willing to allow participation in study Exclusion: None Patient Characteristics: • age range 10 days–5 years 11 months • Grp 1 male 62, female 45; Grp 2 male 103, female 70 • black 66, white 214	Comparisons: 1. Pneumatic otoscopy[a] – unvalidated examiner Dx– : MEE absent Dx+ : MEE present Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Pneumatic otoscopy[b] – unvalidated examiner Dx– : MEE absent Dx+ : MEE present Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Pneumatic otoscopy[c] – unvalidated examiner Dx– : MEE absent Dx+ : MEE present Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 91.37% (127/139) specificity 74.67% (56/75) PPV 86.99% (127/146) NPV 82.35% (56/68) accuracy 85.51% (183/214) prevalence 64.95% (139/214) Comparison 2: sensitivity 98.55% (136/138) specificity 74.67% (56/75) PPV 87.74% (136/155) NPV 96.55% (56/58) accuracy 90.14% (192/213) prevalence 64.79% (138/213) Comparison 3: sensitivity 98.45% (127/129) specificity 81.16% (56/69) PPV 90.71% (127/140) NPV 96.55% (56/58) accuracy 92.42% (183/198) prevalence 65.15% (129/198)

[a] Ears with fluid "suspected" as having effusion are included as false negatives. Fluid-free ears "suspected" as having effusion included as false positives. "Not examined" ear included as false negatives.
[b] Ears with fluid "suspected" as having effusion are included as true positives. Effusion free ears "suspected" as having effusion included as false positives. The one ear "not examined" is not included.
[c] Ears in the "suspect" category are not included. The one ear that was not examined is not included.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
4790 Paradise 1996	Diagnostic study Study Quality Score (0–6): 5 (111110) Examiner(s): • tympanometry (Grason-Stadler GSI 33): two audiologists • pneumatic otoscopy: tested otoscopists which we interpreted as validated otoscopists Grp 1: Ears in infants during their first year of Grp 2: Ears in infants during their second year of life (Note: these ears are not necessarily independent.) N=362 subjects, 1296 ears N1=701 ears N2=595 ears * (did not specify how many children in each group)	Time: not specified Place: Children's Hospital outpatient department Affiliation: Not specified but authors recognized as being from Children's Hospital of Pittsburgh, PA Inclusion: • Age: 2–23 months • healthy infants Exclusion: None Patient Characteristics: Not specified	Comparisons: 1. Grp 1– 1st year of life Examiner type (Audiologist #1) Professional tympanometry Dx– : Normal tympanogram Dx+ : Questionable or Flat Tymp. Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 2. Grp 1– 1st year of life Examiner type (Audiologist #1) Professional tympanometry Dx– : Normal/questionable tymp. Dx+ : Flat tympanograms Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 3. Grp 1– 1st year of life Examiner type (Audiologist #2) Professional tympanometry Dx– : Normal tympanograms Dx+ : Questionable or Flat Tymp. Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 4. Grp 1– 1st year of life Examiner type (Audiologist #2) Professional tympanometry Dx– : Normal/questionable tymp. Dx+ : Flat tympanograms Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 85.71% (120/140) specificity 47.95% (269/561) PPV 29.13% (120/412) NPV 93.08% (269/289) accuracy 55.49% (389/701) prevalence 19.97% (140/701) Comparison 2: sensitivity 47.86% (67/140) specificity 90.91% (510/561) PPV 56.78% (67/118) NPV 87.48% (510/583) accuracy 82.31% (577/701) prevalence 19.97% (140/701) Comparison 3: sensitivity 42.86% (60/140) specificity 90.02% (505/561) PPV 51.72% (60/116) NPV 86.32% (505/585) accuracy 80.60% (565/701) prevalence 19.97% (140/701) Comparison 4: sensitivity 32.14% (45/140) specificity 96.08% (539/561) PPV 67.16% (45/67) NPV 85.02% (539/634) accuracy 83.31% (584/701) prevalence 19.97% (140/701) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 4790 Paradise 1996			*Continued* 5. Grp 2– 2nd year of life Examiner type (Audiologist #1) Professional tympanometry Dx– : Normal tympanograms Dx+ : Questionable/Flat tymp. Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 6. Grp 2– 2nd year of life Examiner type (Audiologist #1) Professional tympanometry Dx– : Normal/questionable tymp. Dx+ : Flat tympanograms Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 7. Grp 2– 2nd year of life Examiner type (Audiologist #2) Professional tympanometry Dx– : Normal tympanograms Dx+ : Questionable/Flat tymp. Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present 8. Grp 2– 2nd year of life Examiner type (Audiologist #2) Professional tympanometry Dx– : Normal/questionable tymp. Dx+ : Flat tympanograms Validated pneumatic otoscopy GS– : effusion absent GS+ : effusion present	*Continued* Unit of measure: Ear Comparison 5: sensitivity 90.36% (75/83) specificity 69.92% (358/512) PPV 32.75% (75/229) NPV 97.81% (358/366) accuracy 72.77% (433/595) prevalence 13.95% (83/595) Comparison 6: sensitivity 65.06% (54/83) specificity 94.92% (486/512) PPV 67.50% (54/80) NPV 94.37% (486/515) accuracy 90.76% (540/595) prevalence 13.95% (83/595) Comparison 7: sensitivity 68.67% (57/83) specificity 92.00% (471/512) PPV 58.16% (57/98) NPV 94.77% (471/497) accuracy 88.74% (528/595) prevalence 13.95% (83/595) Comparison 8: sensitivity 65.06% (54/83) specificity 94.92% (486/512) PPV 67.50% (54/80) NPV 94.37% (486/515) accuracy 90.76% (540/595) prevalence 13.95% (83/595)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
4793 Park 1988	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • tympanometry (Teledyne Avionics instrument type TA-3D and Grason-Stadler instrument 1723 middle ear analyzer): not specified • myringotomy: not specified Study Cohort: Children diagnosed as having OME who received myringotomies with and w/out ventilation tubes *(a group of 79 healthy children were also studied but findings were not relevant to key question four) N=290 subjects, 528 ears	Time: Grp 1~not specified; Grp 2~1981–1986 Place: Korea (specifics not provided) Affiliation: not specified Inclusion: • Grp 1~Age: Not specified • Grp 2~Age: 2–15 years • Grp 1~healthy children, who showed <10 db air-bone gap and in whom physical exams were normal • Grp 2~OME diagnosed by physical exam and impedance audiometry • Korean pediatric population Exclusion: None Patient Characteristics: • Grp 2 age range 2–15 years	Comparisons: 1. Male Children Professional tympanometry Dx– : Type A tympanogram Dx+ : Type B or C tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Male Children Professional tympanometry Dx– : Type A or C tympanograms Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Female Children Professional tympanometry Dx– : Type A tympanogram Dx+ : Type B or C tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 4. Female Children Professional tympanometry Dx– : Type A or C tympanograms Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 91.94% (228/248) specificity 13.16% (5/38) PPV 87.36% (228/261) NPV 20.00% (5/25) accuracy 81.47% (233/286) prevalence 86.71% (248/286) Comparison 2: sensitivity 73.39% (182/248) specificity 39.47% (15/38) PPV 88.78% (182/205) NPV 18.52% (15/81) accuracy 68.88% (197/286) prevalence 86.71% (248/286) Comparison 3: sensitivity 94.82% (183/193) specificity 16.13% (5/31) PPV 87.56% (183/209) NPV 33.33% (5/15) accuracy 83.93% (188/224) prevalence 86.16% (193/224) Comparison 4: sensitivity 78.76% 152/193) specificity 54.84% 17/31) PPV 91.57% (152/166) NPV 29.31% (17/58) accuracy 75.45% (169/224) prevalence 86.16% (193/224) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 4793 Park 1988			*Continued* 5. Male Children Professional tympanometry Dx– : Type A tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 6. Female Children Professional tympanometry Dx– : Type A tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear Comparison 5: sensitivity 90.10% (182/202) specificity 17.86% (5/28) PPV 88.78% (182/205) NPV 20.00% (5/25) accuracy 81.30% (187/230) prevalence 87.83% (202/230) Comparison 6: sensitivity 93.83% (152/162) specificity 26.32% (5/19) PPV 91.57% (152/166) NPV 33.33% (5/15) accuracy 86.74% (157/181) prevalence 89.50% (162/181)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2236 Rees 1992	Cross-sectional Study Quality Score (0–6): 1 (100000) Examiner(s): • tympanometry (Rexton tympanometer): not specified • myringotomy: not specified *This study focuses on changes in tympanometry resulting from general anesthesia, not on the sensitivity/specificity of tympanometry for identifying fluid filled ears. Study Cohort: Children admitted to ENT Dept. for myringtomy and possible grommet insertion N=155 subjects, 310 ears	Time: 6-month period (actual dates not specified) Place: ENT Department of the Radcliffe Infirmary, Oxford, UK Affiliation: Dept. of Otolaryngology, Radcliffe Infirmary, Oxford, UK Inclusion: • Age: 18 months–11 years • admitted to the ENT Dept. of the Radcliffe Infirmary over a 6-month period • seen by an otologist within 3 months of admission • with clinically persistent middle ear effusion • able to tolerate examination and tympanometry Exclusion: None Patient Characteristics: • age range 18 months–11 years	Comparisons: 1. Portable tympanometer[a] Dx– : Type A or C tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Portable tympanometer[b] Dx– : Type A or C tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 100.00% (260/260) specificity 18.00% (9/50) PPV 86.38% (260/301) NPV 100.00% (9/9) accuracy 86.77% (269/310) prevalence 83.87% (260/310) Comparison 2: sensitivity 100.00% (260/260) specificity 100.00% (50/50) PPV 100.00% (260/260) NPV 100.00% (50/50) accuracy 100.00% (310/310) prevalence 83.87% (260/310)

[a] Results are based on the tympanogram readings 1 hour before myringotomy.
[b] Results are based on the tympanograms obtained immediately before myringotomy.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
4804 Renvall 1996	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • tympanometry (Model 85AR [American Electromedics Corporation]): not specified • myringotomy: not specified Study Cohort: Children who underwent tympanometry and were candidates for ventilation tube insertion N=73 subjects, 127 ears	Time: not specified Place: not specified Affiliation: Inclusion: • Age: 3–12 years • suffered from longstanding serous otitis media • candidate for tympanocentesis and ventilation tube insertion under general anesthesia Exclusion: None Patient Characteristics: • age range 3–12 years	Comparisons: 1. Professional tympanometry Dx– : Peaked tympanogram Dx+ : Flat tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present\	Unit of measure: Ear Comparison 1: sensitivity 86.14% (87/101) specificity 96.15% (25/26) PPV 98.86% (87/88) NPV 64.10% (25/39) accuracy 88.19% (112/127) prevalence 79.53% (101/127)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2344 Sassen 1994	Diagnostic study Study Quality Score (0–6): 3 (111000) Examiner(s): • tympanometry (type GSI-27A and type Tymp-85TT, 226 Hz probe tone, used six weeks in each group and then interchanged): not specified • otomicroscopy: three ENT surgeons • myringotomy: three ENT surgeons *This study combines children undergoing myringotomy because of recurrent AOM or chronic OME but does not stratify the results according to type of OM. Grp 1: Hospital A ~ Ears in children with middle ear problems who were only undergoing insertion of tympanostomy tubes [not (adeno)-tonsillectomy with myringotomy] Grp 2: Hospital B ~ Ears in children with URIs who were undergoing (adeno)-tonsillectomy with myringotomy *Continued on next page*	Time: not specified Place: Study performed in two hospital: Hospital A~ Gemini-Hospital in Den Helder and Hospital B~ St. Elisabeth-Hospital in Leiderdorp, The Netherlands Affiliation: Dept. of Otorhinolaryngology, Univ. Hospital Leiden, The Netherlands Inclusion: • Age: 5 months–11 years & 5 months • candidate for insertion of ventilation tubes, or adenoidectomy and/or tonsillectomy w/myringotomy • Grp 1~middle ear problems (not URIs) and undergoing insertion of tympanostomy tubes only • Grp 2~URIs and undergoing (adeno)-tonsillectomy with myringotomy *Continued on next page*	Comparisons: 1. Grp 1 and Grp 2 Binocular micro-tympanoscopy Dx– : Tympanic membrane~normal Dx+ : Tympanic membrane~abnormal Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Grp 1 and Grp 2 Professional tympanometry–both types Dx– : Type (A,C1,C2,O) Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Grp 1 and Grp 2 Professional tympanometry–both types Dx– : Type (A, C1, O) tympanogram Dx+ : Type B or C2) tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 4. Grp 1 and Grp 2 Age (5 months–2 years) Professional tympanometry–both types Dx– : Type (A,C1,C2) tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 96.77% (270/279) specificity 13.86% (14/101) PPV 75.63% (270/357) NPV 60.87% (14/23) accuracy 74.74% (284/380) prevalence 73.42% (279/380) Comparison 2: sensitivity 80.65% (225/279) specificity 58.42% (59/101) PPV 84.27% (225/267) NPV 52.21% (59/113) accuracy 74.74% (284/380) prevalence 73.42% (279/380) Comparison 3: sensitivity 90.68% (253/279) specificity 32.67% (33/101) PPV 78.82% (253/321) NPV 55.93% (33/59) accuracy 75.26% (286/380) prevalence 73.42% (279/380) Comparison 4: sensitivity 90.38% (47/52) specificity 66.67% (10/15) PPV 90.38% (47/52) NPV 66.67% (10/15) accuracy 85.07% (57/67) prevalence 77.61% (52/67) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2344 Sassen 1994	*Continued* N=266 subjects, 515 ears N1=273 ears N2=242 ears * (did not specify how many children in each group)	*Continued* Exclusion: • Grp 1~ URIs • perforation of the tympanic membrane • history of >6 myringotomies and/or >3 insertions of middle ear ventilation tubes Patient Characteristics: • age range 5 months–11 years 5 months • male 143, female 123	*Continued* 5. Grp 1 and Grp 2 Age (5 months–2 years) Professional tympanometry–both types Dx– : Type (A, C1) tympanogram Dx+ : Type (B or C2) tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 6. Grp 1 and Grp 2 Age (2–12 years) Professional tympanometry–both types Dx– : Type (A,C1,C2) tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 7. Grp 1 and Grp 2 Age (2–12 years) Professional tympanometry–both types Dx– : Type (A, C1) tympanogram Dx+ : Type (B or C2) tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 8. Grp 1 Professional tympanometry–both types Dx– : Type (A,C1,C2) tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	*Continued* Unit of measure: Ear Comparison 5: sensitivity 92.31% (48/52) specificity 53.33% (8/15) PPV 87.27% (48/55) NPV 66.67% (8/12) accuracy 83.58% (56/67) prevalence 77.61% (52/67) Comparison 6: sensitivity 81.38% (236/290) specificity 62.60% (82/131) PPV 82.81% (236/285) NPV 60.29% (82/136) accuracy 75.53% (318/421) prevalence 68.88% (290/421) Comparison 7: sensitivity 94.83% (275/290) specificity 35.11% (46/131) PPV 76.39% (275/360) NPV 75.41% (46/61) accuracy 76.01% (320/421) prevalence 68.88% (290/421) Comparison 8: sensitivity 75.42% (135/179) specificity 74.65% (53/71) PPV 88.24% (135/153) NPV 54.64% (53/97) accuracy 75.20% (188/250) prevalence 71.60% (179/250) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2344 Sassen 1994			*Continued* 9. Grp 1 Professional tympanometry–both types Dx– : Type (A, C1, O) tympanogram Dx+ : Type (B or C2) tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 10. Grp 2 Professional tympanometry–both types Dx– : Type (A,C1,C2) tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 11. Grp 2 Professional tympanometry–both types Dx– : Type (A, C1) tympanogram Dx+ : Type (B or C2) tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear Comparison 9: sensitivity 91.62% (164/179) specificity 43.66% (31/71) PPV 80.39% (164/204) NPV 67.39% (31/46) accuracy 78.00% (195/250) prevalence 71.60% (179/250) Comparison 10: sensitivity 90.80% (148/163) specificity 52.00% (39/75) PPV 80.43% (148/184) NPV 70.37% (38/54) accuracy 78.57% (187/238) prevalence 68.49% (163/238) Comparison 11: sensitivity 97.55% (159/163) specificity 30.67% (23/75) PPV 75.36% (159/211) NPV 85.19% (23/27) accuracy 76.47% (182/238) prevalence 68.49% (163/238)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2412 Shaw 1978	Cross-sectional Study Quality Score (0–6): 1 (100000) Examiner(s): • tympanometry (Grason-Stadler 1720 otoadmittance meter, 220 and 660 Hz probe tones): not specified • myringotomy; not specified *This is a pre-, post-anesthesia study looking at the effects on tympanometry. Study Cohort: Ears in children who underwent myringotomy surgery N=59 ears * Sample size =39 patients and 68 ears with probable middle-ear fluid. Only 59 ears received surgery but the authors did not state whether all patients in the sample received surgery or if some were excluded when number of ears was reduced from 68 to 59.	Time: not specified Place: Not specified Affiliation: Univ. of Illinois, Hearing Clinic, Champaign, IL Inclusion: • Age: 2–11 years • probable middle-ear fluid unilaterally or bilaterally based on otologic and audiologic exam • underwent myringotomy surgery Exclusion: • ears that did not undergo myringotomy surgery Patient Characteristics: • age range 2–11 years	Comparisons: 1. Professional tympanometry[a] Dx– : Type A tympanogram Dx+ : Type B and C tympanograms Myringotomy (sedated) GS– : Fluid absent GS+ : Fluid present 2. Professional tympanometry[a] Dx– : Type A and C tympanogram Dx+ : Type B tympanograms Myringotomy (sedated) GS– : Fluid absent GS+ : Fluid present	Unit of measure: Ear Comparison 1: sensitivity 100.00% (49/49) specificity 10.00% (1/10) PPV 84.48% (49/58) NPV 100.00% (1/1) accuracy 84.75% (50/59) prevalence 83.05% (49/59) Comparison 2: sensitivity 97.96% (48/49) specificity 30.00% (3/10) PPV 87.27% (48/55) NPV 75.00% (3/4) accuracy 86.44% (51/59) prevalence 83.05% (49/59)

[a] This is a pre-, post-anesthesia study looking at the effects on tympanometry. Pre-operative results were used.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2545 Szucs 1995	Diagnostic study Study Quality Score (0–6): 1 (100000) Examiner(s): • tympanometry (Amplaid 720 type impedance meter): not specified • myringotomy: ENT surgeon Study Cohort: Children w/ chronic OME and recurrent otitis media N=40 subjects, 78 ears	Time: 6 month period between 5/1992–12/1992 Place: Ear-Nose and Throat Department of Brussels Free University, Brussels, Belgium Affiliation: Dept. of Otorhinolaryngology, Brussels Free Univ., Belgium Inclusion: • Age: 1 year 3 months–11 years 4 months • chronic OME and/or recurrent OM Exclusion: • although not stated specifically to be a criteria none of the children had any accompanying disease Patient Characteristics: • mean age 5 years 8 months, range 1 year 3 months–11 years 4 months • male 27, female 13	Comparisons: 1. Professional tympanometry[a] Dx– : Type A tympanograms Dx+ : Type B tympanograms Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry[b] Dx– : Type A and C tympanograms Dx+ : Type B tympanograms Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 3.Professional tympanometry[c] Dx– : Type A tympanogram Dx+ : Type B and C tympanograms Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 94.74% (36/38) specificity 57.14% (8/14) PPV 85.71% (36/42) NPV 80.00% (8/10) accuracy 84.62% (44/52) prevalence 73.08% (38/52) Comparison 2: sensitivity 73.47% (36/49) specificity 79.31% (23/29) PPV 85.71% (36/42) NPV 63.89% (23/36) accuracy 75.64% (59/78) prevalence 62.82% (49/78) Comparison 3: sensitivity 95.92% (47/49) specificity 27.59% (8/29) PPV 69.12% (47/68) NPV 80.00% (8/10) accuracy 70.51% (55/78) prevalence 62.82% (49/78) *Continued on next page*

[a] The above results table excludes type “C” tympanograms.
[b] The above results table includes type “C” tympanograms as normal.
[c] The above results table includes type “C” tympanograms.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2545 Szucs 1995			*Continued* 4. Professional tympanometry[a] Dx– : Type A and C tympanograms Dx+ : Type B tympanogram Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present 5. Professional tympanometry[b] Dx– : Type A and C tympanograms Dx+ : Type B tympanogram Tympanocentesis (non-sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear Comparison 4: sensitivity 82.61% (19/23) specificity 90.00% (9/10) PPV 95.00% (19/20) NPV 69.23% (9/13) accuracy 84.85% (28/33) prevalence 69.70% (23/33) Comparison 5: sensitivity 57.89% (11/19) specificity 70.59% (12/17) PPV 68.75% (11/16) NPV 60.00% (12/20) accuracy 63.89% (23/36) prevalence 52.78% (19/36)

[a] Adenoid hypertrophy present in this group.
[b] Adenoid hypertrophy absent in this group.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2601 Tom 1994	Diagnostic study Study Quality Score (0–6): 4 (111001) Examiner(s): • tympanometry (Teledyne Screening Meter model no. TA-7A): certified audiologists • myringotomy: surgeon *This study combines data for persistent OME and recurrent AOM but does not stratify the results based on this variable. Study Cohort: Children scheduled to undergo myringotomy with pressure equalization tube insertion for either OME refractory to medical management or recurrent OM N=109 subjects, 213 ears	Time: not specified Place: Not specified Affiliation: Div. of Otolaryngology, Children's Hospital of Philadelphia, PA; Div. of Medical Audiology, Children's Seashore House, PA; Dept. of Otolaryngology, Univ. of Pennsylvania Inclusion: • Age: 5 months–11 years • scheduled to undergo myringotomies w/ pressure equalization tube insertion for either OME refractory to medical management or recurrent OM Exclusion: • ears discovered either at tympanometry or surgery to have small perforations (n=5 ears excluded) Patient Characteristics: • mean age 3 years 8 months, range 5 months–11 years 5 months • male 62, female 47	Comparisons: 1. Professional tympanometry[a] Dx– : Type A tympanogram Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry[b] Dx– : Type A and C tympanograms Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Professional tympanometry[c] Dx– : Type A tympanogram Dx+ : Types B and C tympanograms Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 92.59% (100/108) specificity 51.85% (14/27) PPV 88.50% (100/113) NPV 63.64% (14/22) accuracy 84.44% (114/135) prevalence 80.00% (108/135) Comparison 2: sensitivity 65.36% (100/153) specificity 78.33% (47/60) PPV 88.50% (100/113) NPV 47.00% (47/100) accuracy 69.01% (147/213) prevalence 71.83% (153/213) Comparison 3: sensitivity 94.77% (145/153) specificity 23.33% (14/60) PPV 75.92% (145/191) NPV 63.64% (14/22) accuracy 74.65% (159/213) prevalence 71.83% (153/213)

[a] The above results table excludes all ears with type "C" tympanograms.
[b] The above results table includes type "C" tympanograms as normal.
[c] The above results table includes type "C" tympanograms as abnormal.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2607 Toner 1990	Diagnostic study Study Quality Score (0–6): 2 (100001) Examiner(s): • pneumatic otoscopy: one of the authors • tympanometry (REXTON TYMP82): not specified • myringotomy: not specified Study Cohort: Children undergoing myringotomy due to indication of MEE N=121 subjects, 222 ears	Time: Consecutive 5-month period (actual dates not specified) Place: ENT Department of the Belfast City Hospital, Belfast, UK Affiliation: as above Inclusion: • Age: 18 months–12 years • indication of need for myringotomy (in the majority ~ due to clinically persistent MEE) Exclusion: • pneumatic otoscopy and tympanometry could not be performed • lack of cooperation during test procedure or in allowing removal of excessive wax Patient Characteristics: • age range 18 months–12 years	Comparisons: 1. Pneumatic otoscopy – unvalidated examiner Dx– : mobile Dx+ : immobile Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry[a] Dx– : Type A and C tympanograms Dx+ : Type B tympanogram Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Professional tympanometry[b] Dx– : Type A tympanogram Dx+ : Type B and C tympanograms Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 4. Professional tympanometry[c] Dx– : Type A tympanograms Dx+ : Type B tympanograms Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 87.10% (108/124) specificity 88.78% (87/98) PPV 90.76% (108/119) NPV 84.47% (87/103) accuracy 87.84% (195/222) prevalence 55.86% (124/222) Comparison 2: sensitivity 86.29% (107/124) specificity 92.86% (91/98) PPV 93.86% (107/114) NPV 84.26% (91/108) accuracy 89.19% (198/222) prevalence 55.86% (124/222) Comparison 3: sensitivity 96.77% (120/124) specificity 76.53% (75/98) PPV 83.92% (120/143) NPV 94.94% (75/79) accuracy 87.84% (195/222) prevalence 55.86% (124/222) Comparison 4: sensitivity 96.40% (107/111) specificity 91.46% (75/82) PPV 93.86% (107/114) NPV 94.94% (75/79) accuracy 94.30% (182/193) prevalence 57.51% (111/193)

[a] The above results table includes type “C” tympanograms as normal.
[b] The above results table includes type “C” tympanograms as abnormal.
[c] The above results table excludes all ears with type “C“ tympanograms.

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2675 van Balen 1994	Diagnostic study Study Quality Score (0–6): 4 (111010) Examiner(s): • tympanometry (AR 85 [American Electromedics], 226 Hz probe tone, and microtymp [Welch Allyn], 226 Hz probe tone): general practitioner with special training from ENT department with classification problems discussed with Ear, Nose, and Throat surgeons • myringotomy: Ear, Nose, and Throat surgeon Study Cohort: Children selected by the Ear, Nose and Throat surgeon for uni- or bilateral myringotomy and/or tympanostomy tube insertion N=142 subjects, 284 ears	Time: 9/1990–8/1991 Place: Day-care department, University Children's Hospital, Utrecht and the General Hospital, Overvecht, Utrecht, Netherlands Affiliation: Dept. of General Practice, Univ. of Utrecht, Netherlands Inclusion: • Age: 6 months–12 years • undergoing uni- or bilateral myringtomy and/or tympanostomy tube insertion after referral by General Practitioner Exclusion: • craniofacial malformations and those w/trisomy 21 Patient Characteristics: • age range 6 months–12 years	Comparisons: 1. Portable tympanometer Dx– : Type A and C1 tympanogram curves Dx+ : Type B and C2 tympanogram curves Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 94.23% (147/156) specificity 48.05% (37/77) PPV 78.61% (147/187) NPV 80.43% (37/46) accuracy 78.97% (184/233) prevalence 66.95% (156/233)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2713 Vaughan-Jones 1992	Diagnostic study Study Quality Score (0–6): 2 (100001) Examiner(s): • pneumatic otoscopy: not specified • audiometry (Audioscope [Welch Allyn] and Madsen OB822 audiometer): not specified • tympanometry (Microtymp [Welch Allyn], 226 Hz probe tone, and American Electric Model 85, 256 Hz probe tone): not specified • myringotomy: not specified *For pneumatic otoscopy, the authors do not define a negative or positive test result. Study Cohort: Children undergoing myringotomy N=100 subjects, 200 ears	Time: not specified Place: Audiology Department at Ninewells Hospital, Dundee Affiliation: Dept. of Otolaryngology, Ninewells Hospital, Dundee Inclusion: • admitted with a diagnosis of OME and scheduled to undergo myringotomy Exclusion: • inadequate view of the tympanic membrane or variable audiometric responses Patient Characteristics: • mean age male 6.3 years, female 6.2 years • male 56, female 44	Comparisons: 1. Pneumatic otoscopy – unvalidated examiner Dx– : not defined Dx+ : not defined Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 2. Professional tympanometer Dx– : Type A,C1 & C2 curves Dx+ : Type B curves Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 3. Portable tympanometer Dx– : Type A and C1 curves Dx+ : Type B and C2 curves Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 4. Professional tympanometry Dx– : Type A and C1 tympanograms Dx+ : Type B and C2 tympanograms Myringotomy (sedated) GS– : fluid absent GS+ : fluid present *Continued on next page*	Unit of measure: Ear Comparison 1: sensitivity 89.63% (121/135) specificity 75.38% (49/65) PPV 88.32% (121/137) NPV 77.78% (49/63) accuracy 85.00% (170/200) prevalence 67.50% (135/200) Comparison 2: sensitivity 67.41% (91/135) specificity 93.85% (61/65) PPV 95.79% (91/95) NPV 58.10% (61/105) accuracy 76.00% (152/200) prevalence 67.50% (135/200) Comparison 3: sensitivity 88.89% (120/135) specificity 63.08% (41/65) PPV 83.33% (120/144) NPV 73.21% (41/56) accuracy 80.50% (161/200) prevalence 67.50% (135/200) Comparison 4: sensitivity 88.15% (119/135) specificity 70.77% (46/65) PPV 86.23% (119/138) NPV 74.19% (46/62) accuracy 82.50% (165/200) prevalence 67.50% (135/200) *Continued on next page*

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
Continued 2713 Vaughan-Jones 1992			*Continued* 5. Audiometry – air conduction threshold Dx– : ≤26 dB at 500Hz Dx+ : >26 dB at 500 Hz Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 6. Audiometry – air conduction threshold Dx– : ≤26 dB at 1000Hz Dx+ : >26 dB at 1000 Hz Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 7. Audiometry – air conduction threshold Dx– : ≤26 dB at 2000Hz Dx+ : >26 dB at 2000 Hz Myringotomy (sedated) GS– : fluid absent GS+ : fluid present 8. Audiometry – air conduction threshold Dx– : ≤26 dB at 4000Hz Dx+ : > 26 dB at 4000 Hz Myringotomy (sedated) GS– : fluid absent GS+ : fluid present	*Continued* Unit of measure: Ear Comparison 5: sensitivity 68.15% (92/135) specificity 84.62% (55/65) PPV 90.20% (92/102) NPV 56.12% (55/98) accuracy 73.50% (147/200) prevalence 67.50% (135/200) Comparison 6: sensitivity 59.26% (80/135) specificity 93.85% (61/65) PPV 95.24% (80/84) NPV 52.59% (61/116) accuracy 70.50% (141/200) prevalence 67.50% (135/200) Comparison 7: sensitivity 32.59% (44/135) specificity 95.38% (62/65) PPV 93.62% (44/47) NPV 40.52% (62/153) accuracy 53.00% (106/200) prevalence 67.50% (135/200) Comparison 8: sensitivity 46.67% (63/135) specificity 93.85% (61/65) PPV 94.03% (63/67) NPV 45.86% (61/133) accuracy 62.00% (124/200) prevalence 67.50% (135/200)

Evidence Table 4. Diagnostic Methods (Continued)

Record# Author Year	Study Quality Examiner(s) Group(s) and Sample Size	Time/Place/Affiliation Inclusion/Exclusion Criteria Patient Characteristics	Comparison(s) Influencing Factors Diagnostic Methods (Dx), Cutpoints Gold Standards (GS), Cutpoints	Findings
2758 Watters 1997	Diagnostic study Study Quality Score (0–6): 2 (110000) Examiner(s): • tympanometry (Grayson-Stadler [sic] GS133 Tympanometer): pediatric audiologist • myringotomy: not specified Study Cohort: Children who underwent surgery (grommet insertion) for possible MEE N=501 subjects, 955 ears	Time: 12 month period (11/1/1993–10/1994) Place: Affiliation: Dept's of Otolaryngology and Paediatric Audiology, Radcliffe Infirmary, Oxford, UK Inclusion: • Age: 11 months–15 years • undergoing surgery (insertion of grommets) for possible MEE Exclusion: • children with a 'normal' tympanogram in one or both ears and had surgery cancelled Patient Characteristics: • age range 11 months–15 years	Comparisons: 1. Professional tympanometry[a] Dx– : Type A or C tympanogram Dx+ : Type B tympanogram Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present 2. Professional tympanometry[b] Dx– : Type A tympanogram Dx+ : Type B or C tympanograms Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present 3. Professional tympanometry Dx– : Type A tympanogram Dx+ : Type B tympanogram Myringotomy (sedation unknown) GS– : fluid absent GS+ : fluid present	Unit of measure: Ear Comparison 1: sensitivity 91.14% (679/745) specificity 79.05% (166/210) PPV 93.91% (679/723) NPV 71.55% (166/232) accuracy 88.48% (845/955) prevalence 78.01% (745/955) Comparison 2: sensitivity 99.06% (738/745) specificity 33.81% (71/210) PPV 84.15% (738/877) NPV 91.03% (71/78) accuracy 84.71% (809/955) prevalence 78.01% (745/955) Comparison 3: sensitivity 98.98% (679/686) specificity 61.74% (71/115) PPV 93.91% (679/723) NPV 91.03% (71/78) accuracy 93.63% (750/801) prevalence 85.64% (686/801)

[a] This results table includes type "C" tympanograms as normal.
[b] This results table includes type "C" tympanograms as abnormal.