Otitis Media Middle Ear Effusion Identification and Characterization Using an Optical Coherence Tomography Otoscope

Diego Preciado, MD, PhD1,2, Ryan M. Nolan, MEng, CCRP3, Radhika Joshi, CCRP1,2, Gina M. Krakovsky, APRN1, Anqi Zhang, PhD3, Nickolas A. Pudik3, Nankee K. Kumar2, Ryan L. Shelton, PhD3, Stephen A. Boppart, MD, PhD3, and Nancy M. Bauman, MD1

Abstract

Objective. To determine the feasibility of detecting and differentiating middle ear effusions (MEEs) using an optical coherence tomography (OCT) otoscope.

Study Design. Cross-sectional study.

Setting. US tertiary care children’s hospital.

Subjects and Methods. Seventy pediatric patients undergoing tympanostomy tube placement were preoperatively imaged using an OCT otoscope. A blinded reader quiz was conducted using 24 readers from 4 groups of tiered medical expertise. The primary outcome assessed was reader ability to detect presence/absence of MEE. A secondary outcome assessed was reader ability to differentiate serous vs nonserous MEE.

Results. OCT image data sets were analyzed from 45 of 70 total subjects. Blinded reader analysis of an OCT data subset for detection of MEE resulted in 90.6% accuracy, 90.9% sensitivity, 90.2% specificity, and intra/interreader agreement of 92.9% and 87.1%, respectively. Differentiating MEE type, reader identification of nonserous MEE had 70.8% accuracy, 53.6% sensitivity, 80.1% specificity, and intra/interreader agreement of 82.9% and 75.1%, respectively. Multivariate analysis revealed that age was the strongest predictor of OCT quality. The mean age of subjects with quality OCT was 5.01 years (n = 45), compared to 2.54 years (n = 25) in the remaining subjects imaged (P = .0028). The ability to capture quality images improved over time, from 50% to 69.4% over the study period.

Conclusion. OCT otoscopy shows promise for facilitating accurate MEE detection. The imageability with the prototype device was affected by age, with older children being easier to image, similar to current ear diagnostic technologies.

Keywords
otitis media, acute otitis media, otitis media with effusion, middle ear effusion, optical coherence tomography, otoscopy, imaging, pediatric

Received July 16, 2019; accepted December 26, 2019.

Otitis media (OM) is the most common diagnosis for medical visits in preschool-age children, amounting to more than 25 million visits annually in the United States1 and the most frequent indication for outpatient antibiotic use in the United States, as well as the world, with an estimated US public health cost of at least $2.9 billion annually.2,3 OM is characterized by signs and symptoms of middle ear effusion (MEE) (ie, fluid in the middle ear). Acute otitis media (AOM) is defined by the presence of middle ear inflammation and sudden-onset MEE, often presenting with constitutional symptoms consistent with infection, such as fever and otalgia. Otitis media with effusion (OME) is characterized by MEE without fever, otalgia, or distinct signs of ongoing inflammation that are more typically associated with AOM. OME is overdiagnosed in the primary care setting,4,5 in large part because the tympanic membrane (TM) of a crying child can be mistaken for...
Accurately diagnosing the presence, type, and duration of MEE in pediatric patients is critical to the appropriate management of OM using the current Academy guidelines. Current management of AOM begins with determining the presence or absence of MEE. If MEE is absent, watchful waiting and analgesics can be prescribed. If MEE is present, treatment involves either watchful waiting (eg, serous MEE present) or antibiotic intervention (eg, nonserous MEE present). With recurrent acute infections or with chronic MEE of 3 months or more, tympanostomy tube placement may be beneficial to provide middle ear ventilation, pressure equalization, drainage of MEE, and hearing restoration. However, the presence of MEE in pediatric patients can be difficult to diagnose by simple otoscopy in the office setting, which has been reported as low as 50% accuracy. Failure to detect MEE can have a negative impact on speech and language development in children. In addition to the inherent difficulties when assessing TM surface features, further confounding factors include obstruction of the otoscopy view of the TM due to cerumen and/or narrowing/tortuosity of the ear canal, dim otoscope lighting, and patient noncompliance due to fear of otoscopy. To improve diagnostic accuracy, pneumatic otoscopy and/or tympanometry are recommended by current guidelines. Pneumatic otoscopy aids in the identification of MEE, but this exam technique is reliant on establishing an airtight seal with an ear canal speculum, which can be challenging to accomplish in pediatric patients. In addition, tympanometry can be useful for detecting MEE, but it also requires a seal of the ear canal, as well as significant training to perform correctly. Therefore, pneumatic otoscopy and tympanometry are underused (38% and 7% reported use, respectively), especially in primary care.

A reliable, noninvasive method for detecting and differentiating MEE, without requiring a seal of the ear canal, would be a valuable addition to the routine otoscopic examination. Optical coherence tomography (OCT) is a real-time imaging technique for noninvasively investigating human tissues. Considered the optical analogue of ultrasound imaging, OCT uses a low-intensity light source instead of sound to produce 2-dimensional and 3-dimensional structural images with micron-scale resolution. The image produced by the reflected light is analyzed and can be used to differentiate air from fluid, as well as characterize fluid properties due to scattering of the imaging signal from particles in the fluid. OCT is currently clinically used in ophthalmology as the gold-standard exam for detection of retinal pathology, and there has been rapid growth in many fields of medicine, including cardiology, gastroenterology, dermatology, and oncology. While there is a significant body of research supporting OCT applications in the middle and inner ear, accuracy studies of image readability with an intraoperatively confirmed ground truth for MEE have been limited to a low number of patients. This study therefore aims to evaluate OCT clinical usability and OCT image readability by clinical office personnel. In this study, we explore the feasibility and accuracy of using OCT technology in a modified handheld otoscope to facilitate detection and differentiation of MEE in pediatric patients undergoing tympanostomy tube placement for clinically indicated reasons.

**Materials and Methods**

Children’s National Medical Center Institutional Review Board (IRB) approval was obtained (CNHS 00006493).

Eligible participants included pediatric patients undergoing tympanostomy tube placement under general anesthesia by a pediatric otolaryngologist surgeon (D.P. or N.M.B.) at Children’s National Health System (Washington, DC), regardless of race, ethnicity, or sex. Exclusion criteria included patients for whom standard otoscopy examination was difficult due to sensory issues and patients with stenosed ear canals, including those with trisomy 21. During the preoperative clinic visit, caretakers of eligible participants were introduced to the study, given a pamphlet describing OCT and the study goals, and informed that they would be invited on the day of surgery to participate.

Following written informed consent, and written assent for children age 7 and older, bilateral OCT otoscopy images were collected using a portable OCT otoscope prototype device (PhotoniCare, Champaign, Illinois), complete with a handheld imager with an otoscope form-factor and disposable speculum tips, as shown in Figure 1. The TM was visualized via digital otoscopy displayed on the device monitor. Using a trigger button on the handheld, each button press captured an image set of 10 consecutive 2-dimensional OCT images from the prior 2 seconds’ worth of data displayed, to account for device user reaction delay. Bilateral imaging was attempted (by D.P., R.J., G.M.K., or N.M.B.), with data collected from both ears when possible. Imaging was not completed if (1) there was complete obstruction of the otoscopic TM view due to cerumen or (2) the subject was summoned to the operating room before study images were able to be obtained. Once in the operating room, myringotomy and ventilation tube placement with a binocular surgical microscope was performed per routine standard of care by the surgeon (D.P. or N.M.B.). If MEE was present and aspirated/collected, its appearance was interpreted by the surgeon as serous, purulent, or mucoid based on color, transparency, and perceived viscosity.

Demographic data were collected, deidentified, and stored in a password-protected Research Electronic Data Capture (REDCap; Vanderbilt University, Nashville, Tennessee) database along with the intraoperative MEE findings. Deidentified image data were digitally transferred securely to PhotoniCare for storage and analysis.

OCT otoscope data were sorted into readable (high-quality) and nonreadable (poor-quality) images by a blinded OCT otoscopy expert (R.M.N.) using the following protocol. Images classified as readable required an identifiable
TM signal in the OCT image and a confirmatory view of the TM in the correlated otoscopy image. Images classified as nonreadable had poor or no signal in the OCT image and/or lacked a clear, focused view of the TM in the correlated otoscopy image. Reasons for nonreadable data include (1) the OCT otoscope was not inserted to the correct focal depth (the otoscopy and OCT focal depths were coregistered), (2) the image data were identified as only the ear canal and/or cerumen instead of TM, and (3) the subject became noncompliant during imaging. OCT imageability was defined as the number of subjects with at least 1 ear of readable data per subject enrolled. Age, sex, study period, and surgical location (ambulatory surgical center vs main hospital) were considered variables for multivariate analysis of OCT imageability. Mann-Whitney tests were performed for univariate continuous variable comparisons and Fisher’s exact tests were performed for comparisons of categorical variables (Prism 8; GraphPad Software, San Diego, California). For multivariate analysis, logistic regression was performed with effect sizes reported as odds ratios with corresponding 95% confidence intervals (XLSTAT; Addinsoft, Paris, France).

A blinded OCT reader quiz was created by first sorting 65 ears’ (45 patients) worth of readable data into 3 groups based on the intraoperative MEE findings, then randomly selecting images from within each group to produce a 20-image quiz comprising (1) no fluid (n = 7), (2) serous fluid (n = 6), and (3) nonserous fluid (ie, purulent or mucoid fluid; n = 7). These 20 images were combined with their duplicate mirror images (reflected about the vertical axis of each image) and randomly ordered to complete the 40-image quiz. A brief OCT image training set of 3 representative examples of each fluid classification group was then randomly selected from the remaining images and used to educate readers on how to identify the TM, orient the location of the ear canal and middle ear spaces, and how these groups may look different. Representative images of the full training set are shown in Figure 2. The completed quiz was structured in this way to enable readers to complete the quiz in 20 minutes total (including 5-10 minutes of OCT image training).

Blinded reader analysis of OCT images for identifying presence and type of fluid was compared with intraoperative findings to determine the sensitivity, specificity, accuracy, positive/negative predictive values, and inter/intrareader agreement of OCT otoscopy. Four reader groups (n = 6 each) consisted of (1) otolaryngologists, (2) pediatricians, (3) physician extenders, and (4) nonmedical professionals (ie, business administrators, engineers, technical graduate students). Groups 1 to 3 represent tiered levels of clinical expertise, whereas group 4 is a control group to evaluate for clinical experience dependence. To evaluate for OCT experience dependence, 11 of the 24 total readers had prior experience with ear-specific OCT and 13 had no prior experience. Readers were given a scorecard (Figure 3) and asked to assign a number 0 to 6 to each OCT image. Assignment of a “0” meant a confident “no fluid” case, “3” meant a confident “serous fluid” case, and “6” meant a confident “nonserous fluid” case. To glean some insight into the decision-making process of readers, options 1, 2, 4, and 5 were used to indicate that the readers were less confident with their OCT image interpretation, but “1” indicated their leaning toward “no fluid,” “2” or “4” leaning toward “serous fluid,” and “5” leaning toward “nonserous fluid.” For statistical analysis, scores ≥2 were considered positive for MEE, and ≥5 were considered positive for nonserous MEE.

Results
Over the course of 16 months, 70 participants scheduled for tympanostomy tube surgery were enrolled, with no harms or adverse events encountered. Thirty-nine were male and 31 were female. Mean age was 4.13 years (0.58-17.55 years of age). Forty-three were undergoing tube placement for recurrent AOM and 27 for chronic OM. Readable images were collected in 65 ears from 45 participants. In addition, 42 of 70 (60%) patients were enrolled at the ambulatory surgical center, while 28 of 70 (40%) were enrolled at the main hospital. Overall, OCT imageability was 45 of 70 (64%). The median age of patients in which readable images were obtained was 5.01 years vs 2.54 years for those with nonreadable images (P = .0028). OCT imageability increased over the study period from 50.0% during the first 6 months of the study to 69.8% during the last 6 months (Figure 4). Multivariate analysis controlling for age, sex, study location, and study period revealed a significant association for age (P = .006) and study location (P = .031) to ability to capture...
quality images, with a larger proportion of quality images obtained in older patients and at the ambulatory surgical center (Table 1).

As shown in Table 2, reader detection of MEE had a 90.6% accuracy, 90.9% sensitivity, 90.2% specificity, 94.5% positive predictive value, 84.2% negative predictive value, and intra/interreader agreement of 92.9% and 87.1% (0.724 Fleiss’s κ), respectively. Readers with OCT experience (n = 11) had 91.4% accuracy, 91.3% sensitivity, and 91.6% specificity, and readers without OCT experience (n = 13) had 90.0% accuracy, 90.5% sensitivity, and 89.0% specificity. Assessing the impact of prior OCT experience, by 1-way analysis of variance (ANOVA) (P < .05), there was no statistically significant difference between scores for sensitivity (P = .819), specificity (P = .548), or accuracy (P = .408). Comparing between the 4 reader groups’ accuracy, sensitivity, and specificity, the only statistically significant differences by 1-way ANOVA (P < .05) were for sensitivity

![Figure 2](image1.png)

**Figure 2.** Representative blinded reader analysis training images. (A) Healthy ear images, showing optical coherence tomography cross-sectional view of the tympanic membrane (taken along dotted line). (B) Serous effusion: heterogeneous, low-brightness scattering signal. (C) Nonserous effusion: homogeneous, stronger scattering signal.

![Figure 3](image2.png)

**Figure 3.** Blinded readers scorecard for grading optical coherence tomography images. High diagnostic confidence was indicated by scores of 0 (no fluid), 3 (serous fluid), and 6 (nonserous fluid). Intermediate scores (1 vs 2, 4 vs 5) indicated deliberation.

![Figure 4](image3.png)

**Figure 4.** The percentage of readable, quality optical coherence tomography (OCT) images increased from 50.0% (first 6 months) to 69.4% (last 6 months) from increased user familiarity with the OCT otoscope (not significant on multivariate analysis) and intermittent feedback/training.
for groups 1 vs 2 \((P = .0493)\) and for specificity between groups 1 vs 2 \((P = .00929)\), 1 vs 3 \((P = .00565)\), and 1 vs 4 \((P = .0442)\).

As shown in Table 2, for differentiating MEE type, reader identification of nonserous MEE had a 70.8% accuracy, 53.6% sensitivity, 80.1% specificity, 59.2% positive predictive value, 76.2% negative predictive value, and intra/interreader agreement of 82.9% and 75.1% \((0.424 \text{ Fleiss’s } \kappa)\), respectively. Readers with OCT experience \((n = 11)\) had 73.4% accuracy, 57.8% sensitivity, 81.8% specificity, and 58.0% PPV, and readers without OCT experience \((n = 13)\) had 68.7% accuracy, 50.0% sensitivity, 78.7% specificity, and 55.8% PPV.

### Table 2. Blinded Reader Analysis of Optical Coherence Tomography Images.

<table>
<thead>
<tr>
<th>Source</th>
<th>Odds Ratio</th>
<th>Standard Error</th>
<th>Lower Bound (95%)</th>
<th>Upper Bound (95%)</th>
<th>(P) Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study period</td>
<td>0.174</td>
<td>0.112</td>
<td>–0.050</td>
<td>0.397</td>
<td>.125</td>
</tr>
<tr>
<td>Location</td>
<td>0.246</td>
<td>0.111</td>
<td>0.024</td>
<td>0.468</td>
<td>.031</td>
</tr>
<tr>
<td>Age</td>
<td>0.315</td>
<td>0.112</td>
<td>0.092</td>
<td>0.539</td>
<td>.006</td>
</tr>
<tr>
<td>Sex</td>
<td>–0.018</td>
<td>0.112</td>
<td>–0.240</td>
<td>0.205</td>
<td>.875</td>
</tr>
</tbody>
</table>

*Bold \(P\) values represent significance of \(P < .05\).

### Discussion

This study shows the feasibility of using OCT otoscopy to facilitate detection and differentiation of MEE. Specifically, for the detection of the presence of MEE, the blinded OCT image readers displayed high accuracy, sensitivity, specificity, and predictive values, along with strong inter- and intrareader agreement. These results indicate the potential for OCT to be a viable diagnostic technology that is equally effective in the hands of many device users, regardless of prior experience, with accuracy and sensitivity at least as good as other commercially available diagnostic tools, if not better. Furthermore, the 5 to 10 minutes of OCT image interpretation training enabled the blinded readers to perform well on the quiz and may indicate that brief image interpretation training is sufficient.

As shown in Table 2, blinded OCT reader performance discriminating MEE type between borderline serous and nonserous MEE remains a challenge, due to the need to subjectively stratify the strength or brightness of the OCT signal coming from the MEE. While reader detection of
MEE presence is shown to be fairly straightforward given the rather binary presence or absence of OCT signal in the middle ear space, determination of type of MEE could be further enabled using machine learning algorithms to quantify image differences across varying fluid types. Another limitation of the fluid analysis in this study was that the ground truth for MEE classification in this study was a subjective assessment at the time of surgery into 1 of 3 commonly clinically used categories. For robust algorithm development, future studies must incorporate quantitative MEE laboratory analysis to more accurately and objectively characterize MEE purulence and/or viscosity for correlation with the objective image parameters of OCT otoscopy. Such ability to accurately differentiate between serous and non-serous MEE through objective analysis could lead to more judicious antibiotic stewardship.

Other factors also affected the results of this study. First, on the day of surgery, each subject was examined with a standard video otoscope, followed by the prototype OCT otoscope. In a pediatric subject population, especially in younger children in whom the imageability results were lower, conducting multiple ear exams sequentially within a short time frame can result in decreased subject compliance and therefore data quality. In future studies, the separate initial digital otoscopy exam will not be required. Second, as is true with all video otoscopy, the lack of a viewscreen on the handheld required the user to look away at a monitor while operating the device in the subjects’ ear canals. Yet, quality image capture improved in children over the study duration, believed to be due to improvements in user familiarity with the OCT otoscope device, in part due to intermittent image quality feedback and suggestions for improving performance, although user familiarity was not found to be significant on multivariate analysis. There were different device users at each site, which may also have had an impact on the dependence of imageability on study site. It is also worth noting that in the prototype design and all current ear diagnostic equipment, experienced otoscopy users will perform better than less experienced users. Another limitation of this study is the potential for nitrous oxide at the induction and maintenance of anesthesia to have resolved a middle ear effusion that may have been present preoperatively when imaged.

Identifying and focusing on the appropriate region of the TM was one of the greatest challenges to attaining readable images in this study. Collection of out-of-focus OCT images and imaging the ear canal instead of the TM were common in the beginning of the study since the operators were not accustomed to indirectly viewing the TM on the monitor rather than directly through the handheld, like in traditional otoscopy. In addition, the prototype used in this study had a restricted depth of field compared to a traditional otoscope, which likely resulted in navigation difficulty and misinterpretation of the ear canal as the TM.

A number of improvements to device design would address some of the identified limitations. (1) Placing a digital image display directly on the handheld OCT otoscope will enable more familiarity and would negate the need for the users to pull their attention away from the patient to observe data during imaging. (2) Image analysis algorithms would reduce the amount of training needed and subjectivity required to interpret the data. (3) Providing indicators for OCT image quality directly on the handheld would allow users to solely focus on the patient imaging experience. (4) Ergonomic improvements to the device would improve the familiarity and align the use of the device more with traditional otoscope form factors and grip-style preferences. (5) Optimizing the user interface to minimize challenges and barriers associated with direct interaction with the device would optimize image capture.

Refinements to OCT otoscope image collection should enhance both imageability and usefulness of OCT otoscopy in otolaryngology, pediatrics, emergency, and other departments and clinics where ears are frequently examined for OM. Upcoming studies will assess the generalizability of OCT otoscopy for these providers. OCT otoscopy may also offer important advantages of determining MEE properties and differentiating the type of effusion. Identifying mucoid effusions that are less likely to spontaneously resolve may prove beneficial in restoring hearing sooner to children affected by OME, although the aforementioned improvements will be critical to achieving this goal.

**Conclusion**

OCT otoscopy shows promise as a tool to diagnose the presence or absence of MEE, regardless of OCT or prior medical experience. Development to improve upon the clinical ease of use of this prototype OCT otoscope and facilitate better differentiation of MEE type is needed to address the limitations seen in this study, while striving to provide accurate and efficient utility for device users of different otoscopy experience backgrounds. Future studies will compare the accuracy of OCT otoscopy to current clinical diagnostic technologies and determine its generalizability among health care providers to improve the accurate detection and differentiation of MEE, to drive appropriate management of AOM and OME.

**Author Contributions**

Diego Preciado, study design, data collection, analysis and interpretation of the prototype device image data, and writing of this manuscript; Ryan M. Nolan, study design, prototype device manufacture, data quality review, analysis and interpretation of the prototype device image data, and writing of this manuscript; Radhika Joshi, data collection, analysis and interpretation of the prototype device image data, and writing of this manuscript; Gina M. Krakovsky, data collection, analysis and interpretation of the prototype device image data, and review of this manuscript; Anqi Zhang, prototype device manufacture analysis and interpretation of the prototype device image data, and review of this manuscript; Nickolas A. Pudik, data quality review, analysis and interpretation of the prototype device image data, and writing of this manuscript; Nankee K. Kumar, data collection, analysis and interpretation of the prototype device image data, and review of this manuscript; Ryan L. Shelton, study design, prototype device manufacture,
References

Sponsorships: PhotoniCare personnel manufactured and supplied the imaging device prototypes used in this study, led the study design and conduct, led the analysis and interpretation of the devices’ image data, and contributed to the writing as well as provided final approval of this manuscript.

Funding source: This study was financially supported by the National Institute on Deafness and Other Communication Disorders (NIDCD) of the NIH through a SBIR grant: 1R44DC014599-01.

Disclosures

Competing interests: Diego Preciado, salary support from the National Institutes of Health (NIH) Small Business Innovation Research (SBIR) grant; Ryan M. Nolan, salary support from the NIH SBIR grant, as well as personal and corporate financial ownership interest in PhotoniCare, as a cofounder; Radhika Joshi, salary support from the NIH SBIR grant; Gina M. Krakovsky, salary support from the NIH SBIR grant; Nancy M. Bauman, study design, collection, analysis and interpretation of the prototype device image data, and writing of this manuscript; Stephen A. Boppart, study design, analysis and interpretation of the prototype device image data, and writing of this manuscript; Nancy M. Bauman, study design, data collection, analysis and interpretation of the prototype device image data, and writing of this manuscript.


