No sponsorships or competing interests have been disclosed for this article.

Abstract
Objectives. The inconsistent measures used to report outcomes after mastoidectomy to treat cholesteatoma make it impossible for clinicians to compare results and apply them appropriately to patient care. We sought to identify and assess the type and relative frequency of the reported measures.

Data Sources. PubMed, EMBASE, the Web of Science.

Review Methods. We searched 3 independent databases for articles that reported outcomes of middle ear cholesteatoma treated with mastoidectomy. Articles were assessed for eligibility using Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) protocol and data were extracted by 2 independent investigators. Observational studies and randomized controlled trials from the past 30 years were included. Articles with less than 50 subjects and nonsurgical studies were excluded. The measures used by each study were recorded, and the relative frequency of each measure was calculated.

Results. Forty-three of 380 articles met criteria. Time to follow up was inconsistent. Most articles reported on preoperative qualitative parameters (77%); however, few reported on the severity of disease (19%) or formally staged the disease (12%). Not all studies reported on the presence of recurrent or residual disease (88%) or the postoperative audiogram results (70%). Only 12% presented results as a Kaplan-Meier disease-free curve. Most studies reported on complications (70%).

Conclusion. Reported measures were inconsistent among the studies, which makes comparisons between studies unreliable. Standardization will optimize future reporting and will allow for the establishment of best practices.

Keywords
cholesteatoma, measures, outcomes, mastoidectomy

Introduction
Acquired cholesteatoma is a destructive process that represents the final stages of chronic middle ear disease. It is a disabling disease that can result in numerous complications including intractable aural discharge, profound hearing loss, ossicular destruction, vertigo, facial nerve paralysis, meningitis, and intracranial abscess. Surgery is the mainstay of treatment. The best surgical procedure to treat cholesteatoma has often been debated, and despite decades of cholesteatoma research, the topic remains controversial.1-9 We hypothesize that part of the controversy stems from the lack of standardized reporting. Without standardization, an accurate comparison between studies cannot be achieved and the evidence becomes confusing. It is therefore impossible to assess patients’ postsurgical prognosis and the debate over which surgical technique offers the best outcome continues.

We therefore conducted a systematic review to evaluate the currently reported measures used in the literature for the surgical treatment of middle ear cholesteatoma. We aim to list the reported measures and report on their relative frequency. Our secondary aim is to identify which reported measures are most useful for future investigators to report when designing outcomes studies after mastoidectomy for cholesteatoma.

Methods
We searched 3 independent databases for articles through May 2012 that reported outcomes of middle ear cholesteatoma treated with a mastoidectomy10-12 (Table 1). Adhering to Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting standards articles were assessed by 2 different investigators for their

1Department of Otolaryngology—Head and Neck Surgery, State University of New York Downstate Medical Center, Brooklyn, NY, USA
This article was presented at the New York Metropolitan Resident Research Symposium; May 30, 2014; New York, New York.

Corresponding Author:
Niv Mor, MD, Department of Otolaryngology—Head and Neck Surgery, State University of New York, Downstate Medical Center, 450 Clarkson Avenue, Brooklyn, NY 11203, USA.
Email: nivmor73@gmail.com
eligibility.\textsuperscript{13} Data extraction was performed by 2 independent investigators in order to ensure accuracy. We did not use any non-English databases, unpublished data, or any manual search techniques.

Observational studies and randomized control trials were included in our search. We included only articles from the past 30 years because they were more likely to reflect current surgical management than older articles. We also only included studies with 50 or more subjects since they were more likely to have greater statistical power, use a wide range of measures, and include statistical analysis. Nonsurgical studies were excluded. We intentionally limited our search criteria in the hopes of selecting articles that reported on similar disease processes treated with comparable surgical procedures. We therefore selected only those articles that treated middle ear cholesteatoma and included a mastoidectomy as part of the surgical procedure. The measures used from each study were recorded, and the relative frequency of each measure was calculated.

We performed subgroup analysis to isolate those studies with potentially greater clinical impact by looking at 2 citation metrics (the number of times the article was cited in other journals and the average number of citations per year).\textsuperscript{14} The average for both citation metrics were calculated, and articles that were above the average for both metrics were selected as having a greater clinical impact. The relative frequency of each measure from this subgroup was also calculated.

**Results**

After subtracting the duplicates, our literature search resulted in 380 articles. Fifty-two articles met our inclusion and exclusion criteria. Nine articles that were not available in English were also excluded. Forty-three articles were used in the quantitative analysis.\textsuperscript{1-9,15-48} Figure 1 shows our PRISMA flow diagram.

Articles were published between 1984 and 2010. Authorship came from 13 different countries and 42 different cities. The majority of the studies (90.7%) are retrospective chart review, 7% are prospective cohort studies, 1 study is a randomized controlled trial, and 1 study is a nonrandomized controlled trial. Samples size ranged from 50 to 972 with a median sample size of 108. Reported measures and their relative frequency are reported in Table 2 and Table 3.

Time to follow-up was inconsistent among studies (14% 1 yr or less, 26% >1 yr-5 yrs or less, 35% >5 yrs-10 yrs or less, 23% >10 yrs). Most articles made a preoperative qualitative assessment of disease (77%), which included either the location of the cholesteatoma within the middle ear (65%), the presence of ossicular erosion (42%), or a statement of the severity of disease (19%). Few articles formally staged the disease (12%, n = 5). Some articles stated that the disease was “extensive” without mentioning how this was decided. Of the 5 articles that formally staged the disease, 3 used an established staging system and 2 used their own classification system.\textsuperscript{49-51} (Table 2).

Reported postoperative outcome measures fell into 3 categories, which were also inconsistently reported (presence of a dry ear 47%; integrity of the tympanic membrane 51%; presence of middle ear aeration 16%). Some studies did not report on the occurrence of recurrent or residual disease. Furthermore, the definition of recurrence was not consistent between articles with some articles reporting an isolated retraction pocket as an independent measure from recidivism (28%). Recurrence was reported as a Kaplan-Meier disease-free curve in 12% of the articles (Table 2).

Most, but not all, articles reported their postoperative audiogram results (70%), and there was inconsistency with respect to the timing of the postoperative audiogram (19% 1 yr or less, 16% >1 yrs-5 yrs or less, 7% >5 yrs-10 yrs or less, 0% >10 yrs). In 44% of the articles the author did not state the timing of the postoperative audiogram, and in these situations it was assumed that only the best audiogram result was reported. Fifty-eight percent of the articles mentioned obtaining a preoperative audiogram and less showed the actual preoperative audiogram results (Table 3).

Some articles did not report on the number of procedures needed to achieve their stated results. Specifically, performing additional or unplanned procedures was reported in 61%

### Table 1. Data Sources and Search Terms.

<table>
<thead>
<tr>
<th>Database</th>
<th>Search Terms</th>
<th>Number of Articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMBASE</td>
<td>“cholesteatoma”/exp/dm_su AND “mastoid”/exp/dm_su AND “clinical study”/exp</td>
<td>49</td>
</tr>
<tr>
<td>The Web of Science</td>
<td>“cholesteatoma” AND “mastoid” AND “surgery” AND “follow-up” OR “study”</td>
<td>210</td>
</tr>
</tbody>
</table>

Downloaded from oits.sagepub.com by guest on January 17, 2015
of the articles, and performing a second look procedure that did not require further disease eradication was reported in 58% of articles. Operative complications were reported in 70% of the articles (Table 3).

Among the 43 articles that met our inclusion criteria, 14 articles were above the average in both citation metrics (the average number citations by other journals and the average number of citations per year). They were selected as having a greater clinical impact among the articles. Relative frequency of reported measures from this subgroup did not show increased uniformity (Table 2 and Table 3).

**Discussion**

Our systematic review reveals that there is no uniformity in the literature of the measures used when reporting on the surgical treatment of cholesteatoma. Without standardized reported measures it is difficult to make comparisons between studies, and therefore it is impossible to assess the advantage of any specific surgical procedure.

Strangerup et al emphasized the importance of standardization when they analyzed the rate of recurrence in cholesteatoma. They showed that recurrence rates can differ solely based on the method of statistical calculation. While the focus of this article is not to examine the method of data analysis, the article by Stangerup et al highlights the potential to create artificial disparities with inconsistent reporting. They also emphasized the importance of standardizing the time to follow-up and accounting for those patients who are lost to follow-up. In our study we showed inconsistencies in reported follow-up time, and few articles (12%) reported recurrence as a Kaplan-Meier analyses, which reflects the time to an event and therefore accounts for patients who are lost to follow-up.

Reported hearing outcomes are another area of ambiguous reporting. Excellent hearing results after cholesteatoma surgery is often dependent on patients’ preoperative audiogram. Interpretation of postsurgical audiograms in isolation can potentially be misleading. While it may be fair to assume that all of the authors assessed patients’ hearing before taking them to surgery, only 58% of the articles mentioned obtaining a preoperative audiogram and even fewer studies compared the preoperative audiogram to the audiogram after surgery. Furthermore, we found no consistency with the timing of the postsurgical audiogram, which is important as hearing results may vary the longer patients

---

**Figure 1.** PRISMA flow chart showing our inclusion and exclusion criteria and the number of articles that were selected for each criteria. Abbreviation: PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.
Some studies (44%) did not mention when the audiogram was performed, and it was assumed that the author reported only their best audiogram results. This potential reporting bias would be eliminated if there were a standardized system for reporting outcomes.

Few studies (12%) formally staged the disease, and no universal classification system was used when the disease was staged. One study stated that the disease was “extensive” without describing how this was established. It is not surprising that investigators have difficulty staging the disease. Classification systems are designed to identify cohorts with similar disease-specific survival rates, and the lack of systematic reporting does not allow us to identify which disease parameters have the most significant influence on prognosis. This may explain why numerous cholesteatoma staging systems have been proposed in the literature but none have been universally adopted.49-51,53,54 One may argue that no one classification system has been adopted because middle ear cholesteatoma represents too broad a spectrum of disease, and standardized reporting is one way to validate or dispute this claim.

Most authors agree that the goals of cholesteatoma surgery are the complete removal of disease, achievement of a safe and dry ear, and the preservation and/or restoration of hearing, and we would hope that all articles would at least report on these 3 parameters.2,6,13,16,18,23,26,28-30 However, recurrence was reported in 88% of the articles, the presence of a dry ear was reported in only 47% of the articles, and hearing outcomes were reported in 70% of the articles.

One of our eventual aims of this systematic review was to identify best practices when reporting surgical outcomes for middle ear cholesteatoma. The lack of uniformity not only makes pooling of data for a quantitative meta-analysis impractical, but it makes it difficult to offer recommendations based on reports from the literature alone. Nevertheless, we turn to the list of measures reported in Table 2 and Table 3 as a good starting point. We feel that this list can be used as a template from which to propose and validate best practices. We emphasize the need to systematically report time to follow-up, recurrence rates, hearing outcomes, and complications. The definition of recurrent disease also needs to be clarified. Hearing outcomes should adhere to guidelines put forth by the American Academy of Otolaryngology—Head and Neck Surgery Foundation on reporting hearing outcomes in clinical trials.55 Audiograms should include a scattergram of the patients’ baseline hearing and a postoperative scattergram showing the change in hearing. Outcomes will ideally be reported as a Kaplan-Meier analysis. We recognize that characterizing the severity of disease is difficult as no formal staging system has been adopted. However, the extent of disease should be evaluated and reported before and after surgery. Ideally the preoperative and postoperative measures that assess the severity of the disease will be identical. Future systems to classify disease severity could validate which disease parameters have the most significant influence on prognosis and will validate the utility of reporting on such measures as the location of disease, integrity of the ossicles, presence of a dry ear, integrity of the tympanic membrane, and aeration of the middle ear.

The main limitation of this study is that systematic reviews are limited by the strength of the source articles.

### Table 2. Reported Measures: Follow-up, Assessment of Disease, and Recurrence.

<table>
<thead>
<tr>
<th>Reported Measures</th>
<th>Selected Articles (%)</th>
<th>Greater Clinical Impact (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time to follow up</strong></td>
<td>n = 43</td>
<td>n = 14</td>
</tr>
<tr>
<td>1 yr or less</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>&gt;1 yr-5 yrs or less</td>
<td>26</td>
<td>29</td>
</tr>
<tr>
<td>&gt;5 yrs-10 yrs or less</td>
<td>35</td>
<td>29</td>
</tr>
<tr>
<td>&gt;10 yrs</td>
<td>23</td>
<td>35</td>
</tr>
<tr>
<td>Preoperative assessment of disease</td>
<td>77</td>
<td>71</td>
</tr>
<tr>
<td>Reported on location of cholesteatoma</td>
<td>65</td>
<td>57</td>
</tr>
<tr>
<td>Reported on integrity of ossicular chain</td>
<td>42</td>
<td>43</td>
</tr>
<tr>
<td>Reported on disease severity</td>
<td>19</td>
<td>21</td>
</tr>
<tr>
<td>Reported as stage of disease</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Postoperative assessment of disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reported on dry ear/otorrhea</td>
<td>47</td>
<td>43</td>
</tr>
<tr>
<td>Reported on integrity of TM</td>
<td>51</td>
<td>36</td>
</tr>
<tr>
<td>Reported on aeration of the middle ear</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Recurrence</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reported on recurrent/residual disease</td>
<td>88</td>
<td>93</td>
</tr>
<tr>
<td>Reported RP as a separate outcome</td>
<td>28</td>
<td>36</td>
</tr>
<tr>
<td>Reported as Kaplan-Meier analysis</td>
<td>12</td>
<td>7</td>
</tr>
</tbody>
</table>

Abbreviations: TM, tympanic membrane; RP, retraction pocket.
generated from their search. Our literature search utilized 3 large databases. However, we acknowledge that even the best search engine may miss important publications, and we did not complement our search with other sources such as references listed in primary sources. In addition, we may have excluded some potentially useful articles by restricting our inclusion criteria to English-language articles published in the past 3 decades with a relatively large sample size of at least 50 patients. The great majority of the selected articles came from retrospective studies, which are limited by the reported measures that are available in the chart at the time of the review. This study is also limited by our ability to qualify the impact of any particular article, and we recognize that the relative frequency of citations may not equate with the quality of a study or its true clinical impact within the field.

**Conclusion**

The debate over which surgical procedure to treat middle ear cholesteatoma persists in part due to inconsistencies of reported measures found in the literature. It is possible that no one optimal surgical technique exists and that maybe the extent of disease at the time of diagnosis is the most significant factor in determining patients' prognosis. Standardized reporting can allow accurate comparison between studies and thereby has the potential to uncover which disease or treatment variables are most prognostic. This systematic review underscores the importance of establishing best practices when reporting measures for the surgical treatment of cholesteatoma in the hopes to provide the best clinical care.

**Author Contributions**

Niv Mor, acquisition, analysis, and interpretation of data; drafting the work and revising it; Dina A. Finkel, acquisition, analysis, and interpretation of data; revising the work; Matthew B. Hanson, interpretation of data, revising the work, and final approval of the version to be published; Richard M. Rosenfeld, interpretation of data, revising the work, and final approval of the version to be published.

**Disclosures**

Competing interests: None.

Sponsorships: None.

Funding source: None.

**References**

