Claim Validity of Print Advertisements Found in Otolaryngology Journals

Anthony Del Signore, PharmD, MD; Andrew H. Murr, MD; Lawrence R. Lustig, MD; Michael P. Platt, MD; Scharukh Jalisi, MD; Loring W. Pratt, MD; Jeffrey H. Spiegel, MD

Objective: To evaluate the accuracy and scientific evidence supporting product claims made in print advertisements within otolaryngology journals.

Design: Cross-sectional survey with literature review and multiple-reviewer evaluation. Fifty claims made within 23 unique advertisements found in prominent otolaryngology journals were selected. References to support the claims were provided within the advertisements or obtained through direct request to the manufacturer. Five academic otolaryngologists with distinct training and geographic practice locations reviewed the claims and supporting evidence. Each physician had substantial experience as an editorial reviewer, and several had specific training in research methodology and scientific methods.

Results: Of the 50 claims, only 14 were determined to be based on strong evidence (28%). With regard to the supporting references, 32 references were published sources (76%), while 3 references were package inserts and/or prescribing information (7%). Interobserver agreement among the reviewers overall was poor; however, when 3 or more of the reviewers were in agreement, only 10% of the claims were deemed correct (n=5). Reviewers also noted that only 6% of the claims were considered well supported (n=3).

Conclusion: Advertisers make claims that appear in respectable journals, but greater than half of the claims reviewed were not supported by the provided reference materials.


Author Affiliations:
Departments of Otolaryngology–Head and Neck Surgery, Mt. Sinai Medical Center, New York, New York (Dr Del Signore), University of California San Francisco (Drs Murr and Lustig), and Boston University School of Medicine, Boston, Massachusetts (Drs Platt, Jalisi, and Spiegel). Dr Pratt is in private practice in Fairfield, Maine.
tions. This group is charged with overseeing and approving all materials dispersed by pharmaceutical and device manufacturers. Unfortunately, there is no authority to require companies to submit materials prior to publication, and so unverified information is allowed to circulate to the public while the validation process occurs. Furthermore, according to a 2006 inquiry by the US Government Accountability Office, regulatory letters that were dispersed by the agency were essentially ineffective in preventing misleading information from being published.13

Approximately 70 000 publications advertising drugs and medical devices in all specialties were submitted to the FDA for review in 2007, which was up from 40 000 in 2003.14 This dramatically increased number raises questions whether the FDA can keep pace and thus whether the published claims are valid.15,16 Because so many clinicians rely on information gleaned from advertisements as a type of continuing medical education, and in fact use it to stay abreast of new innovations and therapies,17-20 the information being presented to them must be accurate, reliable, and valid.

To our knowledge, a formal review of the reliability of advertisement information dispersed to otolaryngologists has not been reported in the literature. The paucity of scientific review in this area is unfortunate, given the ubiquity of some otolaryngologic procedures, such as the insertion of tympanostomy tubes and tonsillectomy, which are frequent targets of otolaryngologic advertising.21

METHODS

Claims used in this study were selected from advertisements appearing between January 2007 and July 2008 in Annals of Otolaryngology Rhinology and Laryngology, Archives of Otolaryngology—Head and Neck Surgery, Laryngoscope, and Otolaryngology—Head and Neck Surgery. All drug and medical device advertisements from these journals with 1 or more claims were extracted. A claim was defined as any assertion about product performance that impacted patient care or physician practice. If a specific advertisement appeared more than once among the journals, the duplicate advertisements were not included. If a drug or product had 2 distinct advertisements with separate claims, then both were separately included in the study.

References for the selected claims were noted and compiled. References were defined as materials cited in support of the information or claim in the advertisement. Cited references were categorized as journal articles, data on file (eg, unpublished company documents), meeting abstracts or presentations, and prescribing information (eg, package inserts or Physicians’ Desk Reference entry). For references not identifiable through a literature search using PubMed, 3 attempts were made to contact manufacturers via telephone and/or e-mail to request the supporting data for the advertised claims.

Five board-certified otolaryngologists served as reviewers of the database consisting of the selected advertisements, cited references, and study data sheets. One reviewer was a subspecialty trained otologist/neurotologist; 1 was fellowship trained in head and neck oncology; 1 was fellowship trained in rhinology; and 2 were general otolaryngologists with extensive experience and training. The reviewers were asked to classify claims and rank the level of evidence and appropriateness of cited references. Claims and references were classified into 4 categories according to a modified version of the system created by Loke et al4: (1) unambiguous clinical outcome, (2) vague clinical outcome, (3) emotive or immeasurable outcome, and (4) nonclinical outcome. Reference evidence level was ranked, and an appropriateness score was determined (Figure). Evidence level was defined as 1 of the following: (1) unreferenced; (2) irrelevant reference (readily available but did not support the claim); (3) Nonscientific reference (supported the claim but was not a meta-analysis or randomized control trial [RCT]); (4) limited research–based evidence (included at least 1 supportive and adequate scientific study); and (5) moderate-strong evidence (included at least 1 supportive, relevant, high-quality RCT or multiple supportive, adequate or high-quality studies). Finally, reviewers rated the validity of the claim based on the available material and the confidence that they would use the statement in their practice (Figure).

Weighted k coefficients were calculated for each pair of raters answering each question using Fleiss-Cohen weights thus producing coefficients equivalent to intraclass correlation coefficients. The equivalences across all rater pairs were tested, and the overall weighted k was calculated for each question. When low k scores were obtained, gradings of responses were assessed by a consensus of 3 or more reviewers on a specific score.

RESULTS

Within the 76 journal issues sampled, 370 advertisements were identified. These included a total of 887 claims, of which 367 were referenced (41%). Of the 370 advertisements identified, a total of 77 unique advertisements were included for sampling, which yielded 154 unique claims and 58 associated references for our general pool. Of note, among the 77 advertisements, 50 contained claims, while 27 advertisements served as product placement or awareness advertisements. A range of 1 to 10 claims per advertisement was noted, with an average of 3 claims per advertisement.

Within the 77 unique advertisements, 50 claims were randomly sampled. With regard to subspecialty representation of these claims, head and neck oncology (32%);
n = 16), pediatrics (30% n = 15), and rhinology (28% n = 14) were heavily represented vs otology (10% n = 5) and facial plastic surgery (n = 0). Most of the advertisements sampled dealt with surgical device promotion (58% n = 29) rather than drug promotion (42% n = 21). Within the pool of selected advertisements, 42 unique literature references were cited, of which 76% were to published scientific articles (n = 32), 7% to package inserts (n = 3), 7% to data on file (n = 3), 5% to presentations (n = 2), and 5% to nonscientific articles (n = 2).

Interobserver agreement among reviewers was slight to poor (Table 1). The overall weighted κ is reported for question 4 because no difference existed between weighted κ coefficients across rater pairs. However, pairs differed significantly for questions 3 and 5 (P = .002 and P < .001, respectively), and so overall weighted κ values are not reported. The consensus scores of more than 3 reviewers on an assessment varied by question, with an overall range of 48% (question 3) to 84% (questions 1 and 5).

With regard to the classification of advertisement claims, 56% were found to be unambiguous clinical outcomes (n = 28), while 18% were deemed to be vague outcomes (n = 9). Consensus was unable to be reached for 16% of the claims reviewed (n = 8).

Supporting claims contained some level of research-based evidence in 48% of the advertisements (n = 24), while 18% were nonscientific, irrelevant, or unreferenced claims (n = 9). Reviewers were unable to reach consensus when grading the level of evidence in 34% of the claims (n = 17).

References supported associated claims in 34% of the advertisements reviewed (n = 17), while 12% of the claims actually had references that contradicted part or all of the statement (n = 6). Consensus could not be reached for 52% of the claims reviewed (n = 26).

With respect to the questions regarding accuracy and validity, it was found that 10% of the claims were deemed correct (n = 5), while 50% were not able to be assessed (n = 25). When asked if the accompanying references provided enough support for the claim to be used in daily practice, our panel determined that just 6% of the claims were well supported (n = 3), whereas 58% of the claims were unsupported by the reference materials provided (n = 29).

Table 1. Interobserver Agreement on Claim Evidence Among Reviewers

<table>
<thead>
<tr>
<th>Evidence Characteristic</th>
<th>Weighted κ Range for Reviewer Pairs</th>
<th>Overall Weighted κ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree to which references support the claim (Question 3)</td>
<td>0.064-0.506</td>
<td>NA</td>
</tr>
<tr>
<td>Claim rating based on data provided (Question 4)</td>
<td>0.014-0.386</td>
<td>0.144</td>
</tr>
<tr>
<td>Claim supported enough to use in your clinical practice? (Question 5)</td>
<td>0.072-0.608</td>
<td>NA</td>
</tr>
</tbody>
</table>

Abbreviation: NA, Not applicable.

*For more details on evidence questions evaluated by reviewers, see the Figure.

Our findings within the otolaryngology literature are similar to those found in studies of advertiser claims within other medical and surgical specialties.1-8,15,16 Assurances of data and claim validity by the pharmaceutical industry to professionals and government agencies help to alleviate some of the pressures placed on these firms. In response to previous work raising questions and doubts about advertising, industry officials formed the International Federation of Pharmaceutical Manufacturers and Associations as an attempt to standardize internal and self-review. Membership in the organization requires that manufacturers provide “accurate . . . substantiated . . . and ethical” advertisements.22 Unfortunately, the discordance between reviewers in this study as well as in other well-documented studies illustrates that much work still needs to be done to ensure that valid information is distributed.1-8

According to the Pharmaceutical Research and Manufacturers of America annual industry review, industry trends show that spending on promotions targeting physicians increased from $3.9 billion to $7.2 billion, or about 9% annually from 1997 to 2005.19 Furthermore, a recent study based on publicly available data illustrates that the pharmaceutical industry spends almost twice as much on promotion as it does on research and development.23 Despite this staggering growth in advertising to physicians by pharmaceutical companies, the FDA resources have not increased in proportion to the demands. In the latest budget appropriated by the Office of Management and Budget, the FDA received a 6.7% increase in spending, but the agency’s recent problems have called into question the effectiveness of these allocated resources.29

Given the multitude of concerns that exist about the FDA’s inadequate funding and the pharmaceutical companies’ advertising inaccuracy, others have begun scrutinizing the drug and medical device industry. For example, investigators have shown that 44% to 50% of the statements made in sampled advertisements are poorly supported by the scientific evidence provided.2,7,15,16 In the present study, while 68% of the sampled claims were noted to be supported by moderate to strong research evidence (n = 34), we showed that 58% of the claims sampled were not supported by the provided reference materials (n = 29). In a study by Lankinen et al,7 not a single claim was found to be supported by strong scientific evidence. These data call into question the physician practice of accepting advertiser claims and allowing them to influence patient care5,10,17,20 without investigating the level of the supporting evidence.

A few sampled claims are listed in Table 2. As illustrated by these representative examples, most claims were considered unsupported by our review group. Although we experienced relatively low interobserver agreement among the 5 reviewers, consensus of 3 or more was noted for 84% of the claims (n = 42). It is important to
understand that a reviewer’s assessment of “not supported by the evidence” is not intended to judge the validity of the claim itself. Rather, it means that regardless of the truth of the claim, the evidence presented by the manufacturer is inadequate to support the claim.

With regard to classifying claims, Othman et al25 performed a review of all studies examining the quality of pharmaceutical advertisements using a similar classification scale. Compared with our study in which 50% of the claims were deemed “unambiguous clinical” outcomes, Othman et al25 found that only 25% of the claims fell into that classification. Despite the higher percentage of claims being clearly “clinical” rather than ambiguous or “emotional,” reviewers in our series continued to have difficulty agreeing on the validity of such claims.

Cooper et al26 examined the availability of references to clinicians, comparing advertisement references to scientific article references. Overall, the difference in availability was marked, 84% and 99% for advertisements and scientific research articles, respectively. In other words, while references in scientific publications were unavailable only 1% of the time, the references for supporting articles cited by advertisers were unavailable 16% of the time.26 In addition, although references to journal articles were easily retrieved, other referenced sources (eg, meeting abstracts, presentations, data on file) were not easily acquired. Similarly, we experienced a low rate of receiving requested data on file and presentations despite multiple attempts at contact.

Our study has several limitations that should be noted. One limitation we experienced, as has been noted by other investigators,3 was the relatively low interobserver reliability in reviewer responses. The varied research experiences of our reviewers might account for this, but given their strong qualifications, this is not likely. The reduced interobserver consistency might be due to reviewer opinion regarding the validity of the claim. If an observer believes something to be true, he or she might be more likely to find supporting evidence compelling. It is possible that an evaluation similar to this one could be done using physicians outside of the specialty targeted by the advertisers (ie, nonotolaryngologists). Of course, in this case, a lack of information on the part of the observer might compromise the eventual data, in that the observers will not have the knowledge assumed by the advertiser.

Second, the analysis was limited to 4 journals and examined only a year and a half of publishing. This by no means is representative of all otolaryngologic advertise-ments, but it allows for discussion to be generated, and it highlights some of the inadequacies in the current state of advertising. A more comprehensive analysis using different sampling methods, additional journals, and additional years may have yielded different trends.

Third, although we used a rating scale adapted from a previous study, our low interobserver rating was troubling.11 The low level of agreement further supports the need for clear and transparent advertising to be presented. When well-trained and experienced otolaryngologists cannot agree that the presented data support the advertised claim, we believe that the claims themselves become inherently less convincing. While there will always be controversy and differences of opinion with regard to optimal patient treatment, there should not be disagreement on an objective evaluation of whether a given reference supports a claim.

In the present study, reviewers had full access to the available associated reference materials for each selected claim. Unfortunately, owing to time constraints in the real world of medical practice, readers often blindly accept statements as fact without substantiating the validity and accuracy of the supporting data. This has strong implications that can ultimately jeopardize patient care. The need for manufacturers to make only valid claims and to provide clear supporting data that can be easily interpreted by physicians regardless of research experience is illustrated by these results.

Our study showed that over half of the claims made in print advertisements within the selected sample were not sufficiently supported by the provided reference materials. In addition, disagreement encountered by reviewers evaluating the same pieces of supporting data calls into question the clarity of the supporting documents. With the increasing number of advertisement claims being submitted to regulatory agencies and the dwindling funds appropriated to review these materials, it is imperative that journal editors take a critical look at this important adjunct in patient care.

Submitted for Publication: December 10, 2010; final revision received March 10, 2011; accepted March 29, 2011. Published Online: May 16, 2011. doi:10.1001/archoto.2011.75

Correspondence: Jeffrey H. Spiegel, MD, Department of Otolaryngology—Head and Neck Surgery, Boston University School of Medicine, 830 Harrison Ave, Ste 1400, Boston, MA 02118 (Jeffrey.Spiegel@bmc.org).

Author Contributions: All authors had full access to all

---

Table 2. Examples of Evaluated Claims

<table>
<thead>
<tr>
<th>Claim Type</th>
<th>Claim</th>
<th>Source</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unambiguous</td>
<td>“_________ is proven to relieve a wide range of seasonal allergy rhinitis symptoms . . . with no impairment or sedation.”</td>
<td>Published</td>
<td>Supported</td>
</tr>
<tr>
<td>Vague</td>
<td>“_________ Turbino technique reduces the risk of complications, such as pain, crusting and osteitis.”</td>
<td>Presented at meeting</td>
<td>Unsupported</td>
</tr>
<tr>
<td>Emotive</td>
<td>“_________ Ear Cleansing System generated strong results among chronic ear wax build up sufferers.”</td>
<td>Survey</td>
<td>Unsupported</td>
</tr>
<tr>
<td>Nonclinical</td>
<td>“as compared with electrosurgery or lasers, the ________ allows safer dissection near vital structures.”</td>
<td>Data on File</td>
<td>Unsupported</td>
</tr>
</tbody>
</table>
the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study concept and design: Del Signore and Spiegel. Acquisition of data: Del Signore, Murr, Lustig, Platt, Jalisi, and Spiegel. Analysis and interpretation of data: Del Signore, Lustig, Jalisi, Pratt, and Spiegel. Critical revision of the manuscript: Del Signore, Lustig, and Spiegel. Administrative, technical, and material support: Del Signore, Platt, Pratt, and Spiegel. Study supervision: Jalisi and Spiegel.

Financial Disclosure: None reported.

Previous Presentation: This work was presented as a poster at the Combined Otolaryngology Spring Meeting, May 28-31, 2009; Scottsdale, Arizona.

REFERENCES