Validating the Clinical Assessment of Eustachian Tube Dysfunction: The Eustachian Tube Dysfunction Questionnaire (ETDQ-7)

Edward D. McCoul, MD, MPH, **Vijay K. Anand, MD, FACS**, and **Paul J. Christos, DrPH, MS**
Department of Otolaryngology–Head and Neck Surgery (E.D.M., V.K.A.) and Department of Public Health (P.J.C.), Weill Cornell Medical College, New York, New York, U.S.A.

Abstract

Objectives/Hypothesis—Eustachian tube dysfunction (ETD) is a common condition that is associated with otologic and rhinologic symptoms. The complete assessment of ETD is limited without a valid symptom score. We developed and conducted initial validation of the seven-item Eustachian Tube Dysfunction Questionnaire (ETDQ-7), a disease-specific instrument to assess symptoms with respect to ETD.

Study Design—Validation study.

Methods—The ETDQ-7 was developed using standard survey methodology. The ETDQ-7 was completed by a group of 50 consecutive adult patients diagnosed with ETD and 25 non-ETD patients who served as a control group. Tympanometry was used as a criterion standard to distinguish the two groups. A subset of respondents repeated the ETDQ-7 at a time point 4 weeks later.

Results—Content validity for the ETDQ-7 was established by focus group and review of the literature. Reliability testing indicated acceptable internal consistency for the entire instrument (Cronbach a = .71). The test–retest reliability indicated good correlation between the two questionnaires completed by the same patient 4 weeks apart (r = 0.78). The ETDQ-7 was able to discriminate between patients with ETD and those without (P < .001), indicating excellent discriminant validity.

Conclusions—The ETDQ-7 is a valid and reliable symptom score for use in adult patients with ETD that may facilitate clinical practice by highlighting the impact of ETD. Further testing is needed to determine its usefulness in assessing treatment response.

Keywords

Eustachian tube dysfunction; outcomes research; symptom score; otitis media

INTRODUCTION

Eustachian tube dysfunction (ETD) is a common condition in otolaryngology practice. Although few epidemiologic data exist on the prevalence of ETD, the impact of this

condition can be substantial.[1] Symptoms include aural fullness, tinnitus, a sensation of being "underwater," and an inability to rapidly equilibrate middle ear pressure.[2] Although the burden of ETD on the individual's overall well-being has not been determined, some potential secondary effects of ETD include situational avoidance, communication difficulties, and reduced productivity. A reliable measure of ETD symptoms is necessary for a complete assessment of this condition.

Clinical assessment of symptoms is traditionally a subjective process that includes directed and open-ended questioning, nonverbal cues, and possible clinician bias. Similar to quality-of-life assessments, symptom scores permit the quantitative assessment of subjective domains and have several advantages over the traditional clinical history.[3] First, a symptom score can provide a more precise estimate of disease burden and may yield information not readily identified by the clinician. Second, a symptom score produces formal and validated documentation of patient-reported history for the clinical record. Third, participation in reporting his or her own impressions may motivate patient compliance with prescribed treatment. For all of these purposes, a validated instrument is needed.

Currently, the clinician is limited in the management of ETD by the lack of a validated tool for symptom assessment. To this end, we sought to develop a patient-reported, disease-specific instrument that could be used for the clinical assessment of symptoms and treatment outcome. The present report describes the development and initial validation of this new instrument, the 7-Item Eustachian Tube Dysfunction Questionnaire (ETDQ-7).

MATERIALS AND METHODS

Subjects

Patients were prospectively enrolled from the clinical practice of the senior author (v.k.a.). All subjects were outpatients who presented for otolaryngologic evaluation at a tertiary referral center between August 2010 and October 2010. All patients included in this study were at least 18 years old. Patients were diagnosed as having ETD if they had a retracted or poorly mobile tympanic membrane on pneumatic otoscopy, with a history of at least two of the following symptoms in one or both ears over the previous 1 month period: aural fullness or pressure, a sensation of clogged or muffled hearing, recurrent or persistent middle ear effusion (defined as an effusion present on examinations at least 1 month apart), or the inability to rapidly self-equilibrate middle ear pressure following changes in ambient atmospheric pressure. Abnormal impedance audiometry was used as a criterion standard to verify the diagnosis at the time of enrollment. Exclusion criteria included surgery of the head or neck within 3 months; a history of radiation therapy to the head and neck; sinonasal malignancy; evidence of acute upper respiratory infection, including sinusitis and acute otitis media; adenoid hypertrophy; nasal polyposis; cleft palate or history of cleft palate repair; craniofacial syndrome, including Down syndrome; cystic fibrosis; ciliary dysmotility syndrome; or other systemic immunodeficiency. A second group of patients who did not meet these inclusion criteria and who had presented with medical complaints not related to ETD were consecutively enrolled for use as a control group. Presenting complaints for these patients included voice disturbance, tonsil hypertrophy, and intraoral lesions. All of these patients had a normal examination of the tympanic membrane, middle ear, nasal cavity, and nasopharynx. Normal impedance audiometry was used as a criterion standard to verify the absence of ETD. Written informed consent was obtained from each subject, and approval for this study was obtained from the institutional review board of Weill Cornell Medical College.

Instrument Development

Development of the new instrument began with a review of the literature. Face validity was established by incorporating attributes from existing widely-used, validated instruments in the otolaryngology literature. These included the Otitis Media 6-Item Quality-of-Life Survey (OM-6),[4] the Nasal Obstruction Symptom Evaluation,[5] and the 20-Item Sino-Nasal Outcomes Test (SNOT-20).[6] A discussion group of the authors was convened to enumerate a list of possible questions aimed at assessing symptoms in individuals who presented with ETD. Item relevance was established by consensus of the authors, and five items were included in this preliminary instrument, plus an inquiry of laterality. To provide a focused critique of item relevance, this preliminary version was distributed to a focus group of 10 patients from the senior author's practice that had been diagnosed with ETD. Each respondent was asked to answer each question and to comment on clarity, comprehensibility, and ease of use. Each respondent was also asked to identify any items that were confusing or ambiguous and to write in any additional symptoms that he or she had experienced related to ETD that were not in the questionnaire.

Based on the responses from the focus group, several changes were made to the instrument. Three additional questions were added, and one of the existing questions was divided into two separate items. The qualitative response fields were replaced with a seven-item Likert scale, with a response of "1" indicating no problem and "7" indicating a severe problem. The overall layout was modeled on the layout of other popular questionnaires, notably the OM-6 and SNOT-20. The resulting instrument included nine items plus an inquiry of laterality. Scoring was possible in one of two ways. The total item score could be reported, with a range from 9 to 63, although this lacked the intuitive property of round numbers. Alternatively, the score could be reported as a mean item score, and expressed as a range from 1.0 to 7.0. The second method was preferred because of the easily understood score limits.

This nine-item instrument was then subjected to a pilot run in a sample of 15 patients in the senior author's practice who had complaints of ETD. These responses were tabulated and subjected to reliability testing to establish overall internal consistency and internal consistency with each item deleted. The Cronbach a coefficient for the entire nine-item instrument used in this 15-person sample was .93, indicating a good internal consistency of the instrument for all respondents.

Instrument Validation

Following the development process, the preliminary instrument was formally distributed for data collection and validation testing. This nine-item instrument was administered on an outpatient basis to 50 consecutive patients presenting to the senior author's practice with a diagnosis of ETD. During this period, the same instrument was administered to a sample of 25 consecutive patients who had presented for complaints unrelated to ETD who were to be used as a control group. With a sample size of 50 ETD patients, a two-sided 95% confidence interval for the intraclass correlation coefficient (i.e., Cronbach a) could be constructed to be within approximately ±12.7% of the true value. This calculation assumed a Cronbach a of 70% at the beginning of the study. Sample size calculation was performed in nQuery Advisor Version 7.0 (Statistical Solutions, Cork, Ireland).[7] All patients with ETD who declined initial treatment and returned for follow-up 4 weeks later in an untreated state were asked to complete the instrument a second time to provide data for analysis of test–retest reliability.

Internal consistency reliability was assessed by calculating Cronbach a for the entire instrument and Cronbach a for the instrument with each item deleted. Internal consistency

was considered adequate if Cronbach a was ≥70. Items that did not have adequate correlations were considered to detract from the overall reliability of the instrument and were subsequently marked for deletion. Two questions from the nine-item instrument were eliminated owing to a negative impact on overall internal consistency. Analysis of the remaining questions by comparison of corrected item-total correlations determined that there were no statistically relevant subgroupings within the total question set. The final seven-item instrument was named the ETDQ-7 (Table I).

Construct validity was determined by analyzing content validity and discriminant validity. Content validity was established in the development process outlined; specific strategies included literature review, focus group analysis, preliminary response testing, and focused content revision. Discriminant validity was assessed by comparing the item responses of patients with ETD to those who did not have ETD. The two-sample t test was used to compare ETDQ-7 score between these two groups. Further testing of discriminant validity was performed using receiver operating characteristic (ROC) analysis of the ETDQ-7 responses. Estimates of sensitivity and specificity at the optimal score cut-point, area under the curve (AUC), and an ROC curve were calculated. Test–retest reliability was assessed by calculation of the Spearman rank correlation coefficient.

The assessment of concurrent validity was limited, as there is currently no validated instrument for the assessment of ETD symptoms or disease impact. Many patients with nasal complaints such as rhinosinusitis and nasal obstruction also note symptoms referable to ETD; therefore, we speculated that a disease-specific instrument that assesses sinonasal health might give information that complements and possibly overlaps with the ETDQ-7. The 22-item Sinonasal Outcome Test (SNOT-22) was selected because of its ready availability, ease of administration, and multiple validations in the published literature.[8-10] The Spearman rank correlation coefficient was computed to assess the degree of linear correlation between the ETDQ-7 and SNOT-22. ROC analysis was performed to evaluate the discriminant validity of the SNOT-22 for assessing ETD in comparison to the ETDQ-7.

All P values are two-sided with statistical significance evaluated at the .05 a level. All analyses were performed using SPSS Version 19.0 (SPSS Inc., Chicago, IL).

RESULTS

Subjects

A total of 75 patients were enrolled for the validation portion of the study. Fifty subjects carried a diagnosis of ETD as defined, and 25 had complaints not related to ETD and served as a control group. Patients in the ETD and control groups had a mean (standard deviation) age of 49.8 (13.9) years and 53.4 (11.2) years, respectively. The diseased group contained 24 (48%) men and 26 (52%) women, and the control group contained 13 (52%) men and 12 (48%) women.

Criterion Validity

Impedance audiometry was regarded as the most reliable objective measure of ETD and was accepted as the criterion standard to establish external validity. All subjects meeting inclusion criteria for ETD were found to have an abnormal tympanogram. Conversely, all subjects included in the non-ETD control group were found to have normal tympanometry.

Internal Consistency Reliability

Internal consistency reliability testing of the ETDQ-7 yielded a Cronbach a of .711 for the entire instrument (95% confidence interval, 0.570-0.818). An evaluation of internal

consistency with each item deleted did not materially improve the observed internal consistency (Table II). This demonstrated that all questions were close to measuring the same underlying construct, and as a result no additional items were removed from the instrument.

Test–Retest Reliability

Fifteen patients with ETD returned for follow-up in an untreated state 4 weeks after the initial visit, when they completed the ETDQ-7 a second time. The test–retest reliability for this group was good (Spearman rank correlation coefficient = 0.78), indicating correlation between the two questionnaires completed by the same patient over time. No medical or surgical intervention was prescribed during the intervening period.

Discriminant Validity

Overall ETDQ-7 score among the 50 patients in the ETD group was significantly greater than the score among the 25 patients in the control group ($t = 12.2$, $P < .001$). The mean (standard deviation) overall score was 4.0 (1.1) for the ETD group and 1.3 (0.3) for the control group. The mean individual score for each of the seven items of the ETDQ-7 was significantly greater for the diseased group compared to the control group (Table III).

ROC analysis supported the excellent discriminate validity of the ETDQ-7 (Fig. 1). Using an optimal total item score cutpoint of ≥14.5 versus <14.5 (for categorizing a patient as having ETD) provided 100% sensitivity and 100% specificity. This equated to an ETDQ-7 mean item score of ≥2.1 to indicate the presence of ETD.

Concurrent Validity

Assessment of concurrent validity was limited as there is no preexisting standard for the assessment of ETD symptoms. Responses on the ETDQ-7 showed fair correlation with responses on the SNOT-22 (Spearman rank correlation coefficient = 0.63). In particular, higher scores on the ETDQ-7 were seen in association with higher scores on the SNOT-22, indicating a moderate degree of agreement with an established, validated measure of sinonasal morbidity. ROC analysis using an optimal SNOT-22 score cutpoint of ≥28.5 versus <28.5 (for categorizing a patient as having ETD) provided 84% sensitivity and 72% specificity (Fig. 2). The mean SNOT-22 score was higher for diseased patients as compared to control patients (43.0 ± 16.9 vs. 24.0 ± 9.2, respectively; $P < .0001$, by t test).

DISCUSSION

The current work describes the development and initial validation of the ETDQ-7, a new instrument for the assessment of ETD-related symptoms. We believe that this is the first disease-specific instrument for ETD to be validated in the literature. We have reported normative cross-sectional data for untreated adults with ETD, which can be applied to patients in the clinical setting. The ETDQ-7 provides information that can help to focus the clinical encounter and has potential utility for outcomes research in the study of patients with ETD.

The need for a validated, disease-specific instrument for ETD is particularly notable because of the lack of a widely accepted objective measure of the presence and severity of this disorder. Several objective measures have been proposed, including impedance audiometry,[11] otoscopic appearance,[12] visual grading of endoscopic findings,[13] and invasive tubomanometry.[14] However, at the present time an ideal modality has not been identified.[15] The availability of a valid symptom score is likely to aid the clinician in recording an accurate description of the untreated condition.

The ETDQ-7 is brief and easy to use, and the respondent burden is minimal. It has been shown to be reliable and valid for the cross-sectional assessment of ETD-related symptoms in adults. In particular, the ability of the ETDQ-7 to discriminate between diseased and non-diseased groups was excellent. Retesting of untreated patients with the ETDQ-7 at separate time points showed good test–retest reliability. The ETDQ-7 showed excellent sensitivity and specificity. Comparison with the SNOT-22 showed that although the SNOT-22 has moderate specificity and sensitivity for detecting ETD, the ETDQ-7 performs significantly better as a disease-specific instrument. Criterion validity was established by the exclusive presence of normal tympanograms in control subjects and abnormal tympanograms in subjects with ETD.

Notably, the ETDQ-7 is not intended for the assessment of eustachian tube symptoms that arise in conjunction with acute upper respiratory infection or a neoplastic process, as these patients were excluded from the study group. All patients had chronic symptoms of ETD present for at least 1 month and had not received medical or surgical treatment for their condition. These criteria help to prevent the inclusion of patients with acute or self-limiting disease, which may have an underlying etiology that differs from chronic ETD.

Some limitations of the ETDQ-7 bear mention. First, the response items are primarily concerned with severity of disease burden. The timing of events—in particular, whether symptoms are intermittent or continual or worse during a particular time of day—is not represented in the majority of items. Second, a larger cohort may be helpful to improve the precision of the confidence interval around the Cronbach a. However, with respect to discriminant validity, the analysis was not underpowered because a difference was positively detected between the two groups (i.e., no false negative). Third, because of limited follow-up by the study group, only a subset was available for calculating test–retest reliability. Fourth, the absence of an existing instrument for the assessment of ETD precludes the establishment of concurrent validity. Nonetheless, the correlation of higher SNOT-22 scores with higher ETDQ-7 suggests that sinonasal symptoms may have a causative effect on ETD. In addition, question 4 ("ear symptoms when you have a cold or sinusitis") necessarily overlaps with one possible symptom of rhinosinusitis; however, although some patients experience ear symptoms when they have an upper respiratory infection, many other patients do not. As with the other items that appear in the final instrument, this question was not eliminated because internal consistency testing showed that it contributed to the overall strength of the instrument. Finally, the optimal recall period for symptom assessment in ETD has not been determined. The recall period for the ETDQ-7 was arbitrarily set at 1 month, although a different recall period may have resulted in different overall responses. The recall period inherently represents a compromise, because a longer period improves the sensitivity for detecting positive symptoms but also increases the likelihood of recall bias.

Disease-specific instruments can serve as important outcome measures for clinical interventions. Useful attributes that contribute to validity for outcome measurement include responsiveness, sensitivity to clinical change, and criterion validity. A prospective study of these aspects of the ETDQ-7 is warranted to determine its utility for outcome assessment after the medical or surgical treatment of ETD.

CONCLUSION

The ETDQ-7 is a valid and reliable disease-specific symptom score for adult patients with ETD. It is easily administered in the clinical setting and has minimal burden to respondents. A standardized symptom score may enhance clinical care by highlighting the impact of ETD

and guiding appropriate management. Further prospective testing of patients being treated for ETD may establish the utility of the ETDQ-7 in the assessment of treatment outcomes.

Acknowledgments

BIBLIOGRAPHY

1. Bluestone, CD. Eustachian Tube: Structure, Function, Role in Otitis Media. BC Decker; Hamilton, Ontario: 2005.

2. Seibert JW, Danner CJ. Eustachian tube function and the middle ear. Otolaryngol Clin North Am. 2006; 39:1221–1235. [PubMed: 17097443]

3. Van Wijk R. Assessment of quality of life: advantages and pitfalls. Clin Exp Allergy Rev. 2005; 5:32–35.

4. Rosenfeld RM, Goldsmith AJ, Tetlus L, Balzano A. Quality of life for children with otitis media. Arch Otolaryngol Head Neck Surg. 1997; 123:1049–1054. [PubMed: 9339979]

5. Stewart MG, Witsell DL, Smith TL, Weaver EM, Yueh B, Hannley MT. Development and validation of the Nasal Obstruction Symptom Evaluation (NOSE) Scale. Otolaryngol Head Neck Surg. 2004; 130:157–163. [PubMed: 14990910]

6. Piccirillo JF, Merritt MG, Richards ML. Psychometric and clinimetric validity of the Sino-Nasal Outcome Test (SNOT-20). Otolaryngol Head Neck Surg. 2002; 126:41–47. [PubMed: 11821764]

7. Elashoff, JD. nQuery Advisor Version 7.0. Cork. Statistical Solutions; Ireland: 2007.

8. Morley AD, Sharp HR. A review of sinonasal outcome scoring systems–which is best? Clin Otolaryngol. 2006; 31:103–109. [PubMed: 16620328]

9. Hopkins C, Gillett S, Slack R, Lund VJ, Browne JP. Psychometric validity of the 22-item Sinonasal Outcome Test. Clin Otolaryngol. 2009; 34:447–454. [PubMed: 19793277]

10. Schalek P, Otruba L, Hahn A. Quality of life in patients with chronic rhinosinusitis: a validation of the Czech version of SNOT-22 questionnaire. Eur Arch Otorhinolaryngol. 2010; 167:473–475. [PubMed: 20033700]

11. Poe DS, Silvola J, Pyykko I. Balloon dilation of the cartilaginous Eustachian tube. Otolaryngol Head Neck Surg. 2011; 144:563–569. [PubMed: 21493236]

12. Kujawski OB, Poe DS. Laser Eustachian tuboplasty. Otol Neurotol. 2004; 25:1–8. [PubMed: 14724483]

13. Poe DS, Metson RB, Kujawski O. Laser Eustachian tuboplasty: a preliminary report. Laryngoscope. 2003; 113:583–591. [PubMed: 12671411]

14. Ockermann T, Reineke U, Upile T, Ebmeyer J, Sudhoff HH. Balloon dilatation Eustachian tuboplasty: a clinical study. Laryngoscope. 2010; 120:1411–1416. [PubMed: 20564474]

15. McCoul ED, Lucente FE, Anand VK. Evolution of Eustachian tube surgery. Laryngoscope. 2011; 121:661–666. [PubMed: 21305559]
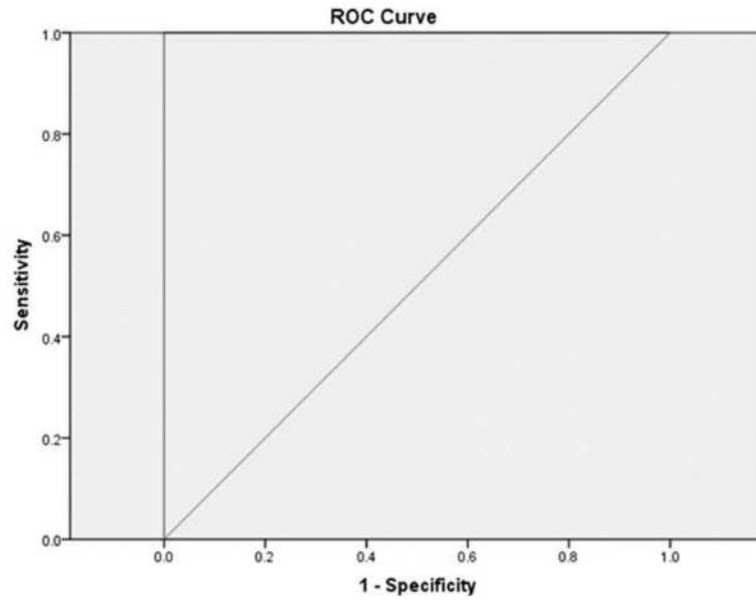
Fig. 1.
Receiver operating characteristic (ROC) curve for the seven-item Eustachian Tube Dysfunction Questionnaire in detecting eustachian tube dysfunction. Area under the curve = 100% (P <.0001).
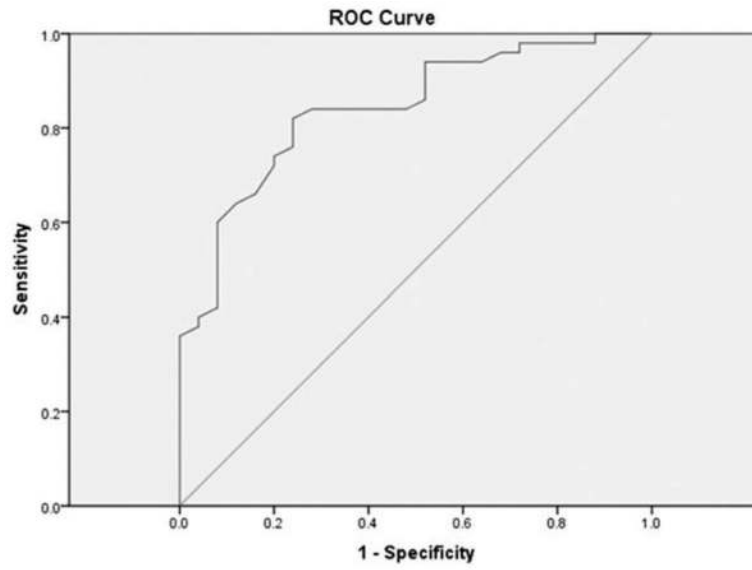
Fig. 2.
Receiver operating characteristic (ROC) curve for the 22-Item Sinonasal Outcome Test in detecting eustachian tube dysfunction. Area under the curve = 84% (*P* < .0001).

TABLE I

The Seven-Item Eustachian Tube Dysfunction Questionnaire.

Over the past 1 month, how much has each of the following been a problem for you?	No Problem		Moderate Problem			Severe Problem	
1. Pressure in the ears?	1	2	3	4	5	6	7
2. Pain in the ears?	1	2	3	4	5	6	7
3. A feeling that your ears are clogged or "under water"?	1	2	3	4	5	6	7
4. Ear symptoms when you have a cold or sinusitis?	1	2	3	4	5	6	7
5. Crackling or popping sounds in the ears?	1	2	3	4	5	6	7
6. Ringing in the ears?	1	2	3	4	5	6	7
7. A feeling that your hearing is muffled?	1	2	3	4	5	6	7

TABLE II

Reliability Testing of Eustachian Tube Dysfunction Questionnaire Showing Overall Internal Consistency and Internal Consistency With Each Item Deleted.

Item	Cronbach α if Item Deleted
1. Pressure	.676
2. Pain	.652
3. Feeling clogged	.668
4. Cold/sinusitis problems	.676
5. Crackling/popping	.693
6. Ringing	.704
7. Feeling muffled	.678
Overall	.711

TABLE III

Eustachian Tube Dysfunction Questionnaire Scores by Item.

Item	ETD Group, Mean (SD)	Control Group, Mean (SD)	_P_ Value[*]
1. Pressure	4.3 (1.5)	1.1 (0.4)	<.001
2. Pain	3.2 (1.7)	1.3 (0.6)	<.001
3. Feeling clogged	4.8 (1.5)	1.2 (0.5)	<.001
4. Cold/sinusitis problems	5.0 (1.6)	2.2 (1.1)	<.001
5. Crackling or popping	4.0 (1.8)	1.1 (0.4)	<.001
6. Ringing	3.2 (2.1)	1.1 (0.3)	<.001
7. Feeling muffled	4.8 (1.5)	1.1 (0.4)	<.001
Overall	4.0 (1.1)	1.3 (0.3)	<.001

ETD = eustachian tube dysfunction; SD = standard deviation.

[*] Paired _t_ test.