Detection of Arytenoid Dislocation Using Pixel-valued Cuneiform Movement

***Ashley P. O'Connell Ferster, †Michael C. Ferster II, ‡Haley Glatthorn, ‡Bartholomew J. Bacak, and †Robert T. Sataloff,** **Hershey and †‡Philadelphia, Pennsylvania*

Summary: Objectives. This study aims to assess utility of pixel-valued movement software in detecting arytenoid dislocation preoperatively.

Study Design. This is a retrospective analysis.

Methods. Twenty-seven patients diagnosed with unilateral arytenoid dislocation were included. Diagnosis of arytenoid dislocation was confirmed by lack of vocal fold paralysis on preoperative laryngeal electromyography and by intraoperative findings of cricoarytenoid dislocation. A region-tracking software algorithm developed by Zhuang et al was used to analyze 27 preoperative endoscopic videos of patients diagnosed with arytenoid dislocation. Vector analysis measuring cuneiform movement during inspiration was used as an indirect measure of arytenoid movement. Values were normalized using vocal fold length. Two raters blinded to diagnosis of arytenoid dislocation measured vocal fold length and cuneiform movement on both the dislocated and the nondislocated sides.

Results. A Wilcoxon signed-rank test indicated that the mean pixel-valued cuneiform movement and standard deviation (SD) were greater for nondislocated (159.24, SD = 73.35) than for dislocated (92.49, SD = 72.11) arytenoids (Z = 3.29, $P = 0.001$). The interrater correlation coefficient was 0.87 for the dislocated side and 0.75 for the nondislocated side. The intrarater correlation coefficient was 0.87 for the dislocated side and 0.91 for the nondislocated side. The receiver operating characteristic curve revealed an area under the curve between 0.76 and 0.83 (95% confidence interval 0.63–0.90). Analysis by the first and second raters revealed misdiagnosis of laterality of arytenoid dislocation in four and six patients, respectively.

Conclusions. The software program developed by Zhuang et al provides a high-degree of precision, with good interrater and intrarater correlation coefficients. However, high rates of misdiagnosis of arytenoid dislocation and the laborious analysis process using this software program make it of limited utility as a clinical diagnostic tool in its present state.

Key Words: Arytenoid dislocation–Vocal fold paralysis–Pixel-valued cuneiform movement–Videolaryngoscopy–Endoscopy.

INTRODUCTION

Arytenoid dislocation (AD) is rare, most commonly occurring secondary to traumatic intubation.[1] Although some authors have suggested that AD occurs in nearly 0.1% of surgeries requiring intubation, those authors suggest that this figure may underestimate the true frequency of AD.[2,3] AD is commonly misdiagnosed as recurrent laryngeal nerve (RLN) paralysis because both conditions classically present with persistent hoarseness, dysphonia, and dysphagia, and decreased vocal fold mobility on laryngoscopy.[4] However, it is critical to differentiate AD from RLN because the former is best treated early with surgical intervention to reduce the dislocated arytenoid.[5]

During normal inspiration, the arytenoid cartilages (ACs) and adjacent cuneiform tubercles abduct from the midline.[6] Breathing typically causes cyclical abduction and adduction of the ACs in a lateral, rotational trajectory that can be tracked from the anterior commissure, which remains relatively stationary.[6] AD compromises movement of the cricoarytenoid joint, resulting in reduced abduction and adduction.[7] AD is frequently interchanged with the term arytenoid subluxation, which is a partial (incomplete) dislocation that alters the relationship between the two cartilages. Although AD is actually a more complete disruption of the joint, these terms are used interchangeably both clinically and within this manuscript.

For patients with RLN paralysis, movement of the cricoarytenoid joint is compromised as well.[7] The nuances of arytenoid movement can be difficult to identify with laryngoscopy alone and, even with stroboscopy, it may be challenging to differentiate AD from RLN paralysis. When AD is suspected, laryngeal electromyography (LEMG) is often used to rule out RLN paralysis, which causes fibrillation potentials, decreased recruitment, and sometimes positive sharp waves.[8] Conversely, LEMG activity in patients with AD shows an absence of these findings, although paresis (typically not paralysis) and AD can co-exist.[8] Although some authors have noted that LEMG has some limitations, it is a safe, effective, and widely recommended routine assessment for vocal fold motion abnormalities.[6–10]

Recently, Zhuang et al developed and tested custom software to track arytenoid movement in videostroboscopic recordings to objectively differentiate AD and RLN paralysis.[11] In Zhuang et al's study, eight patients with AD were included, as well as five patients with RLN paralysis. The authors found that the reduction in arytenoid movement was significantly greater in RLN paralysis than in AD.[11] Their result suggests that

Accepted for publication December 19, 2017.

From the *Department of Surgery, Division of Otolaryngology—Head & Neck Surgery, Penn State Health: Milton S. Hershey Medical Center, Hershey, Pennsylvania; †Department of Otolaryngology—Head & Neck Surgery, Drexel University College of Medicine, Philadelphia, Pennsylvania; and the ‡Drexel University College of Medicine, Philadelphia, Pennsylvania.

Address correspondence and reprint requests to Ashley P. O'Connell Ferster, Otolaryngology—Head & Neck Surgery, Penn State Health: Milton S. Hershey Medical Center, 500 University Drive, Hershey, PA 17033. E-mail: aoconnell2@hmc.psu.edu

Journal of Voice, Vol. 33, No. 3, pp. 370–374
0892-1997

https://doi.org/10.1016/j.jvoice.2017.12.009

computerized tracking of arytenoid movement may be used to diagnose AD.

To further investigate the utility of this software, we tested the method of Zhuang et al on a larger group of patients diagnosed with unilateral AD using their nondislocated sides as controls. We hypothesized that our results would depict greater displacement of cuneiform tubercles in controls than in sides with AD, thus supporting the utility of this software for objective diagnosis of AD.

MATERIALS AND METHODS

The study was approved by the Medical Institutional Review Board at Drexel University College of Medicine in Philadelphia, Pennsylvania. Patients in a tertiary care laryngology practice diagnosed with unilateral AD in the last 20 years were identified. Patient charts were reviewed to confirm diagnosis of unilateral AD after preoperative LEMG of the cricothyroid, thyroarytenoid (TA), and posterior cricoarytenoid (PCA) muscles and intraoperative confirmation of AD. Patients were excluded if they had findings of RLN paralysis on preoperative LEMG or absence of AD during examination in the operating room. All LEMG and operative findings were documented for each patient. Demographic data were recorded, including patient age, etiology of AD, and side or direction of AD.

A total of 38 patients had videostroboscopic examinations available for analysis. Videostroboscopic examinations were uploaded and analyzed with Zhuang et al's custom *MATLAB* (The MathWorks, Inc., Natick, MA) software program.[11] Each video was reviewed to identify a single inspiratory event following phonation of /i/. The video was trimmed to include this sequence of events, thereby showing the transition from minimum vocal fold abduction to maximum glottal opening. Frame sequences in which the anterior commissure or cuneiform tubercles were not visible were not included in the analysis. If no acceptable video frame sequence was identified, the patient was excluded from the analysis. This resulted in exclusion of 10 patients. For the purposes of statistical analysis, one patient also was excluded due to having bilateral AD.

Based on the previous criteria, a total of 27 patients were included in the study. The position of the anterior commissure was identified, tracked, and considered a stable point in the videostrobe frames. Cuneiform tubercles were tracked to follow arytenoid movement in a manner consistent with Zhuang et al.[11] Cuneiform tubercle movement was followed from the point of minimal arytenoid abduction to maximal abduction as seen during normal inspiration. This movement was followed after selecting the cuneiform tubercle and the anterior commissure as the points to be tracked throughout the video frames. The software program then performed its coded tasks, calculating the starting and ending values for cuneiform tubercle position (CT_x, CT_y) and anterior commissure position (AC_x, AC_y). Vocal fold length (V) was also measured by selecting the anterior- and posterior-most limits of the true vocal fold during inspiration.

To account for lateral (ie, translational) changes of the endoscopic camera in a horizontal plane due to patient and examiner movement, the pixel-valued displacement of the anterior commissure was calculated between the first and the last frames and subtracted from the cuneiform tubercle displacement. This produced a vector that represented cuneiform tubercle displacement independent of translational motion. The pixel-valued movement for each cuneiform tubercle was then computed using

$$CT_{x,y} = \sqrt{\left(\Delta CT_X - \Delta AC_X\right)^2 + \left(\Delta CT_Y - \Delta AC_Y\right)^2},$$

where $CT_{x,y}$ is the size in pixels of the path length traveled, ΔCT is the change in cuneiform tubercle position in pixels, and ΔAC is the change in anterior commissure position in pixels.

When visualizing the vocal tract through videostroboscopy, the larynx appears larger when the camera is close and smaller when the camera is far from the larynx, according to a proportionality constant.[11] Additionally, patients with a long vocal tract typically experience longer movement paths than those with a short vocal tract. In accordance with Zhuang et al,[11] movement vectors were normalized using vocal fold length as a scaling factor. The ratio of cuneiform tubercle movement to vocal fold length was then calculated, giving a unitless measurement:

$$CT_{unitless} = CT_{x,y} \,/\, V,$$

where $CT_{unitless}$ is the unitless path length traveled, $CT_{x,y}$ is the size in pixels of the path length traveled, and V is the vocal fold length in pixels. Values were measured by two of the authors (A.O.F., M.F.), with rater 1 (A.O.F.) being a novice user of *MATLAB* and rater 2 (M.F.) having extensive *MATLAB* experience and familiarity with software development. The user's experience is important to assess whether the level of experience impacts the quality of each user's results. Measurements were obtained for the side of the AD and the nondislocated arytenoid.

Statistical analysis was performed with commercial software (*IBM SPSS Statistics* for Windows, Version 24, IBM Corp., Armonk, NY). A Wilcoxon signed-rank sums test was used to compare paired measurements of dislocated and nondislocated arytenoids. Sensitivity analysis was performed using receiver operating characteristic curves. Statistical significance was designated for P values < 0.05. Intraclass and interclass correlation coefficients were calculated to determine intrarater and interrater reliability. Intrarater and interrater reliability were assessed based on accepted rating scales, with an intraclass or interclass correlation coefficient being "excellent" if greater than 0.90, "good" if between 0.75 and 0.90, "fair" if 0.50–0.75, and "poor" if less than 0.50.

RESULTS

Twenty-four patients had LEMG performed before arytenoid reduction procedures. Of these patients, 3 had normal results on LEMG, whereas 5 had bilateral paresis and 16 had paresis on the side of the AD (Table 1). All included patients had findings of AD confirmed on intraoperative examination. Twenty-six patients had dislocation resulting from intubation or extubation trauma; one patient had AD as a result of blunt neck trauma.

The custom software program was used to assess movement of the cuneiform turbercle for both the dislocated and the nondislocated sides. The mean pixel-valued movement of the cuneiform tubercle was determined for each side, along with

TABLE 1.
Summary of Etiology of Dislocation, Side of Dislocation, and Preoperative Laryngeal Electromyography (LEMG) Results for Patients With Arytenoid Dislocation

Patient Number	Side of Dislocation	Results of LEMG
1	Left	Normal
2	Left	Left SLN 30%–40% recruitment
3	Left	Left RLN 70%–80%% recruitment
4	Right	Right RLN 30% recruitment
5	Left	N/A
6	Right	Right RLN and SLN 70% recruitment
7	Left	Normal
8	Left	N/A
9	Left	Left SLN 30% recruitment, left RLN 20% recruitment
10*	Right	Right RLN 50% recruitment
11	Left	Left SLN 30% recruitment, left RLN 70%–80% recruitment
12	Left	Left SLN 70%–80% recruitment, left RLN 30%–40% recruitment
13	Left	N/A
14	Right	Right SLN 70% recruitment
15	Left	Left SLN 70%–80% recruitment, left RLN 60%–70% recruitment
16	Left	Left RLN 30% recruitment
17	Right	Right SLN 70% recruitment, right RLN 60%–70% recruitment
18	Left	Left SLN 50% recruitment, left RLN 40–50% recruitment; right SLN and RLN 70%–80% recruitment
19	Left	Normal
20	Right	Right SLN 40% recruitment, right RLN 70%–80% recruitment
21	Left	Left SLN 80% recruitment, left RLN (TA 60%–70%, PCA 70%) recruitment; right SLN 80% recruitment
22	Left	Left SLN 70% recruitment, left RLN 30%–40% recruitment; right SLN 70%–80% recruitment, Right RLN 40%–50% recruitment
23	Right	Right SLN 40%–50% recruitment; left SLN 85–90% recruitment
24	Right	Right SLN 75% recruitment, right RLN 80%–90% recruitment; left SLN 75% recruitment, left RLN 80%–90% recruitment
25	Left	Left SLN 90% recruitment, left RLN 40%–50% recruitment; right SLN 90% recruitment
26	Right	Right SLN 80% recruitment
27	Left	Left SLN and RLN 40%–50% recruitment

* Etiology of arytenoid dislocation was blunt neck trauma; all other patients' arytenoid dislocation caused by intubation or extubation trauma.
Abbreviations: SLN, superior laryngeal nerve; RLN, recurrent laryngeal nerve (includes PCA and TA muscles unless otherwise noted); PCA, posterior cricoarytenoid muscle; TA, thyroarytenoid muscle.

the unitless, normalized values (Tables 2 and 3). Associated standard deviation and *P* values were calculated as well.

Of the patients examined by rater 1 (A.O.F.), 6 of 27 patients (22.2%) were found to have greater cuneiform tubercle displacement on the dislocated side than the nondislocated side, indicating a misdiagnosis using the software. Of those patients examined by rater 2 (M.F.), 4 of 27 patients (14.8%) were improperly diagnosed with AD.

Intrarater and interrater intraclass correlation coefficient were determined based on the pixel-valued, unitless, and vocal fold length measurements (Table 4). Good to fair reliability was noted based on ratings of intraclass correlation coefficient values. Sensitivity analysis was performed using receiver operating characteristic curves (Figure 1). Between the two raters' results, we calculated areas under the curve ranging from 0.76 to 0.83, which demonstrates fair to good diagnostic strength.

TABLE 2.
Pixel-valued Movement of the Cuneiform Tubercle in Patients With and Without Arytenoid Dislocation*

	Mean Pixel-valued Movement (SD)	Z	P
Dislocated	92.46 (68.50)	3.29	0.001
Nondislocated	159.53 (81.51)		

* Statistical significance was set at *P* value < 0.05.
Abbreviations: SD, standard deviation; Z, Wilcoxon-signed rank test.

TABLE 3.
Unitless Values of the Cuneiform Tubercle in Patients With and Without Arytenoid Dislocation*

	Mean Unitless Values (SD)	Z	P
Dislocated	0.23 (0.14)	4.25	0.0002
Nondislocated	0.43 (0.21)		

* Statistical significance was set at *P* value < 0.05.
Abbreviations: SD, standard deviation; Z, Wilcoxon-signed rank test.

TABLE 4.
Intrarater and Interrater ICCs for Pixel-valued, Unitless, and Vocal Fold Length Measurements Based on Analysis by Raters 1 and 2

ICC (1-1 vs. 2)	ICC	Mean	SD
Vocal fold length	0.978		
CT pixels (nondislocated)	0.765		
CT pixels (dislocated)	0.901		
CT unitless (nondislocated)	0.747	0.439	0.209
CT unitless (dislocated)	0.780	0.232	0.142
ICC (1-1 vs. 1-2)	**ICC**	**Mean**	**SD**
Vocal fold length	0.993		
CT pixels (nondislocated)	0.917		
CT pixels (dislocated)	0.996		
CT unitless (nondislocated)	0.908	0.414	0.209
CT unitless (dislocated)	0.869	0.235	0.143

Abbreviations: 1-1, first value by first rater 1; 1-2, second value by first rater 1; 2, value by second rater; CT, cuneiform tubercle; ICC, intraclass correlation coefficient; SD, standard deviation.

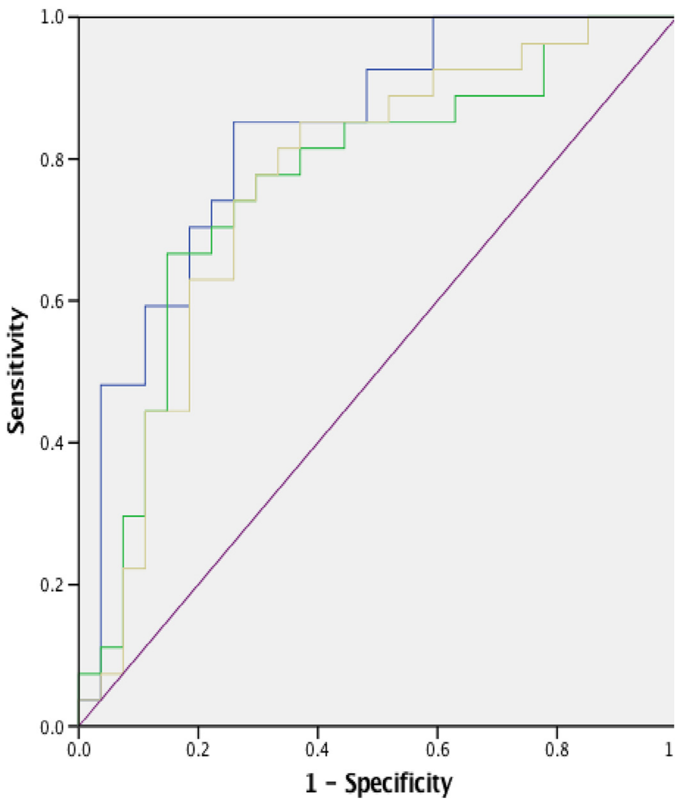

FIGURE 1. Sensitivity analysis using receiver operating characteristic curve with areas under the curve ranging from 0.763 to 0.834. The *purple*, *green*, and *gold curves* correspond to raters 1 and 2 (first and second trials), respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

DISCUSSION

Accurate diagnosis of AD is a challenge for clinicians. The software used in this study was developed by Zhuang et al, in an attempt to diagnose AD objectively and prevent the need for an intraoperative or in-office diagnostic examination by palpation. Their study reported encouraging results, offering a potential adjunct to the clinician's armamentarium when encountering a patient with possible AD.[11]

Values calculated from Zhuang et al's study showed the mean pixel-valued cuneiform movement for AD to be 81.16 (standard deviation [SD] = 25.62) and unitless values to be 0.58 (SD = 0.17).[11] These were compared with values calculated for patients with vocal fold paralysis, which had been confirmed by LEMG. Our study attempted to expand upon their preliminary work by using the software in a larger population. In place of the vocal fold paralysis group, we used retrospective knowledge of sidedness of AD to identify two study groups: arytenoid dislocation and nondislocation.

Our study group had presence and laterality of AD confirmed with both LEMG and intraoperative findings. For our patients, LEMG showed that both TA and PCA muscles were affected similarly, with only patient 21 demonstrating a minor difference in recruitment values (Table 1). Other studies have shown that the TA muscle is more commonly affected, which may be due to intramuscular hemorrhage due to the laryngeal injury causing the AD.[1,10,12] For most of our patients, decreased recruitment was mild and not sufficient to explain the vocal fold immobility as seen on laryngoscopy. Paresis may also be evidence of traumatic neuropraxia following laryngeal injury, although, without preinjury LEMG, pre-existent paresis cannot be ruled out.[13,14] Therefore, with such LEMG findings, one must consider AD as the etiology for patients presenting with impaired vocal fold mobility.

Using the software to compare the unilateral arytenoid dislocation with the nondislocated arytenoid cartilage, the authors simulated a clinical experience in which the software would be needed to detect both presence and laterality of dislocation. Our analysis found dislocation values similar to those reported in the Zhuang et al paper, although our data showed much wider range of values, as noted by our large SDs (Tables 2 and 3). Misdiagnosis could be due to the larger SDs in our study population. Analyses by both of our study's raters showed misidentification of the side of the AD by the software program's analysis. A total of seven of the study's patients were misdiagnosed with respect to laterality; three of the same patients were misdiagnosed by both raters' analyses. This raises concern about misdiagnosis and improper treatment should this software be used in its current form in a clinical setting.

Several flaws inherent to the software analysis also raised concerns about the program's utility in practice. Although the study normalizes the cuneiform tubercle displacement to the vocal fold length, camera rotation and intra-sequence camera movement in a vertical plane are not accounted for in the calculation. Thus, changes in scale of the tracked laryngeal landmarks due to camera motion could not be avoided unless the laryngoscope (rigid laryngeal telescope) was fixed in space relative to the patient's vocal tract.

Variations in lighting also greatly affected the ability to analyze the videos. Slight alterations in lighting were noted even within individual video sequences, often resulting in the tracker falling away from the originally selected anatomic point. Reflections of light on the laryngeal mucosal surfaces also caused the tracker to track inaccurately the anatomic landmarks that were selected, as these reflections moved with movement of the larynx.

The software also was quite difficult to use. For both raters, the use of this program was quite laborious. Video analysis took 1–2 hours to perform, and repeated attempts were often necessary to obtain proper tracking of the selected anatomic point. During each video analysis, the software required input of a number of parameters, which were heuristic in nature. This included a neighborhood size, which was given a standard value of 3 pixels to follow the values used by Zhuang et al, and a template size, which, in our study, ranged from 15 to 45 pixels and was altered based on the software's ability to track the points of interest throughout the video frames. The ability to track these points varied between videos and could not be standardized in our study nor Zhuang et al's study. Parameter values often were edited with repeated tracking attempts, as needed for the software to function with each individual video. Therefore, parameter values were often provided on a "guess and check" basis to produce results. Anatomic points that resembled a corner (eg, anterior commissure) or had distinct color differences along a surface were tracked more easily, whereas more homogeneous surfaces were not tracked well with the software program.

At this time, the interesting software program developed by Zhuang et al would benefit from additional modification to improve accuracy in diagnosis of AD. This is essential to the software's future clinical utility. The concept developed by Zhuang et al is creative and attempts to address an important need. Additional studies are warranted, as well as modifications to the program for this technology to become a clinically useful method to detect AD objectively.

CONCLUSIONS

The software program developed by Zhuang et al provides a high degree of precision, with good interrater and intrarater correlation coefficients. However, high rates of misdiagnosis of AD and the laborious analysis process using this software program make it of limited utility as a clinical diagnostic tool in its present state.

Acknowledgments

We would like to thank Drexel University for their support of this research.

REFERENCES

1. Sataloff RT, Bough ID, Spiegel JR. Arytenoid dislocation. *Laryngoscope*. 1994;104:1353–1361.
2. Goz V, Qureshi S, Hecht A. Arytenoid dislocation as a cause of prolonged hoarseness after cervical discectomy and fusion. *Global Spine J*. 2012;3:47–50.
3. Leelamanit V, Sinkijcharoenchai W. A promising new technique for closed reduction of arytenoid dislocation. *J Laryngol Otol*. 2012;126:168–174.
4. Saigusa H, Kokawa T, Aino I, et al. Arytenoid dislocation: a new diagnostic and treatment approach. *J Nippon Med Sch*. 2003;70:382–383.
5. Sulica L. The natural history of idiopathic unilateral vocal fold paralysis: evidence and problems. *Laryngoscope*. 2008;118:1303–1307.
6. Lindestad PA, Hertegard S, Bjorck G. Laryngeal adduction asymmetries in normal speaking subjects. *Logoped Phoniatr Vocol*. 2004;29:128–134.
7. Rubin AD, Sataloff RT. Vocal fold paresis and paralysis. *Otolaryngol Clin North Am*. 2007;40:1109–1131.
8. Xu W, Han DM, Hou LZ, et al. Laryngeal electromyographic characteristics of vocal fold immobility. *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi*. 2006;41:653–656.
9. Sulica L, Carey B, Branski RC. A novel technique for clinical assessment of laryngeal nerve conduction: normal and abnormal results. *Laryngoscope*. 2013;123:2202–2208.
10. Sataloff RT, Mandel S, Manon-Espaillat R, et al., eds. *Laryngeal Electromyography*. 3rd ed. San Diego, CA: Plural Publishing, Inc; 2017:1–223.
11. Zhuang P, Nemcek S, Surender K, et al. Differentiating arytenoid dislocation and recurrent laryngeal nerve paralysis by arytenoid movement in laryngoscopic video. *Otolaryngol Head Neck Surg*. 2013;149:451–456.
12. Yin SS, Qui WW, Stucker FJ. Value of electromyography in differential diagnosis of laryngeal joint injuries after intubation. *Ann Otol Rhinol Laryngol*. 1996;105:446–451.
13. Rubin AD, Hawkshaw MJ, Moyer CA, et al. Arytenoid cartilage dislocation: a 20-year experience. *J Voice*. 2005;19:687–701.
14. Norris BK, Schweinfurth JM. Arytenoid dislocation: an analysis of contemporary literature. *Laryngoscope*. 2010;121:142–146.