

# Consensus-Based Attributes for Identifying Patients With Spasmodic Dysphonia and Other Voice Disorders

Christy L. Ludlow, PhD; Rickie Domangue, PhD; Dinesh Sharma, PhD; H. A. Jinnah, MD, PhD; Joel S. Perlmutter, MD; Gerald Berke, MD, PhD; Christine Sapienza, PhD; Marshall E. Smith, MD; Joel H. Blumin, MD; Carrie E. Kalata, MS; Karen Blindauer, MD; Michael Johns, MD; Edie Hapner, PhD; Archie Harmon, PhD; Randal Paniello, MD; Charles H. Adler, MD, PhD; Lisa Crujido, MS; David G. Lott, MD; Stephen F. Bansberg, MD; Nicholas Barone, PhD; Teresa Drulia, PhD; Glenn Stebbins, PhD

**IMPORTANCE** A roadblock for research on adductor spasmodic dysphonia (ADSD), abductor SD (ABSD), voice tremor (VT), and muscular tension dysphonia (MTD) is the lack of criteria for selecting patients with these disorders.

**OBJECTIVE** To determine the agreement among experts not using standard guidelines to classify patients with ABSD, ADSD, VT, and MTD, and develop expert consensus attributes for classifying patients for research.

**DESIGN, SETTING AND PARTICIPANTS** From 2011 to 2016, a multicenter observational study examined agreement among blinded experts when classifying patients with ADSD, ABSD, VT or MTD (first study). Subsequently, a 4-stage Delphi method study used reiterative stages of review by an expert panel and 46 community experts to develop consensus on attributes to be used for classifying patients with the 4 disorders (second study). The study used a convenience sample of 178 patients clinically diagnosed with ADSD, ABSD, VT MTD, vocal fold paresis/paralysis, psychogenic voice disorders, or hypophonia secondary to Parkinson disease. Participants were aged 18 years or older, without laryngeal structural disease or surgery for ADSD and underwent speech and nasolaryngoscopy video recordings following a standard protocol.

**EXPOSURES** Speech and nasolaryngoscopy video recordings following a standard protocol.

**MAIN OUTCOMES AND MEASURES** Specialists at 4 sites classified 178 patients into 11 categories. Four international experts independently classified 75 patients using the same categories without guidelines after viewing speech and nasolaryngoscopy video recordings. Each member from the 4 sites also classified 50 patients from other sites after viewing video clips of voice/laryngeal tasks. Interrater  $\kappa$  less than 0.40 indicated poor classification agreement among rater pairs and across recruiting sites. Consequently, a Delphi panel of 13 experts identified and ranked speech and laryngeal movement attributes for classifying ADSD, ABSD, VT, and MTD, which were reviewed by 46 community specialists. Based on the median attribute rankings, a final attribute list was created for each disorder.

**RESULTS** When classifying patients without guidelines, raters differed in their classification distributions (likelihood ratio,  $\chi^2 = 107.66$ ), had poor interrater agreement, and poor agreement with site categories. For 11 categories, the highest agreement was 34%, with no  $\kappa$  values greater than 0.26. In external rater pairs, the highest  $\kappa$  was 0.23 and the highest agreement was 38.5%. Using 6 categories, the highest percent agreement was 73.3% and the highest  $\kappa$  was 0.40. The Delphi method yielded 18 attributes for classifying disorders from speech and nasolaryngoscopic examinations.

**CONCLUSIONS AND RELEVANCE** Specialists without guidelines had poor agreement when classifying patients for research, leading to a Delphi-based development of the Spasmodic Dysphonia Attributes Inventory for classifying patients with ADSD, ABSD, VT, and MTD for research.

*JAMA Otolaryngol Head Neck Surg.* 2018;144(8):657-665. doi:10.1001/jamaoto.2018.0644  
Published online June 21, 2018.

← Invited Commentary  
page 666

+ Supplemental content

**Author Affiliations:** Author affiliations are listed at the end of this article.

**Corresponding Author:** Christy L. Ludlow, PhD, Department of Communication Sciences and Disorders, James Madison University, PO Box 449, Washington, VA 22747 (christyll@gmail.com).

Spasmodic dysphonia (SD) is a rare task-specific disorder<sup>1</sup> affecting voice production during speech.<sup>2</sup> Signs include voice breaks owing to spasmodic vocal fold hyperadduction during vowels in the adductor (ADSD) type<sup>3,4</sup> and prolonged voiceless consonants owing to spasmodic hyperabduction prior to vowels in the abductor (ABSD) type.<sup>5</sup> No diagnostic test has been developed for SD and clinical diagnosis depends on specialists in voice care teams including otolaryngologists, speech-language pathologists (SLP), and neurologists experienced in the identification and treatment of SD,<sup>6</sup> voice tremor (VT),<sup>7</sup> and muscular tension dysphonia (MTD) considered a functional voice disorder.<sup>8</sup> Many patients have limited access to specialized voice centers and often see multiple physicians before diagnosis, thereby delaying treatment by an average of 4.4 years.<sup>9</sup>

The lack of a method for classification of SD type, not only leads to delays in diagnosis and treatment,<sup>10</sup> but impedes research for studying these disorders.<sup>11</sup> Although experienced members in a voice care team agree on classification of patients with SD,<sup>12-15</sup> voice care teams from different centers may not agree. Visualization of the larynx by laryngoscopy is needed to exclude other laryngeal abnormalities that interfere with voice production such as vocal fold paralysis (VFP) or laryngeal lesions. Excluding other laryngeal disorders is required before identifying SD,<sup>10,11,13,16</sup> VT,<sup>13</sup> or MTD.<sup>8,17</sup> Although these disorders have similarities, the optimal treatment differs for each. The symptoms of ADSD are reduced by thyroarytenoid muscle botulinum toxin injections at regular intervals<sup>18,19</sup> or selective adductor denervation with reinnervation surgery.<sup>20,21</sup> Voice tremor is less responsive to botulinum toxin injections than ADSD,<sup>7,22</sup> and MTD commonly responds to voice therapy.<sup>23,24</sup>

Measures conducted in SD include perceptual voice ratings,<sup>25</sup> acoustic measures,<sup>4,5</sup> and laryngeal electromyography.<sup>26</sup> None have demonstrated validity for differentiating SD from other voice disorders. Approaches suggested for differentiating ADSD from MTD include identifying phonatory breaks in speech,<sup>14</sup> using sentence-level material,<sup>27</sup> and analysis of high-speed vocal fold vibration regularity.<sup>28</sup>

We conducted 2 studies. First, we measured experts' agreement on SD classification without using standardized guidelines when they reviewed speech and nasolaryngoscopic video recordings of unfamiliar patients. Second, voice experts employed the Delphi Method<sup>29-31</sup> with 4 reiterative stages to reach consensus on attributes for classifying patients.<sup>32-37</sup> Our aim was to develop classification methods to support research on ADSD, ABSD, VT, and MTD.

## Methods

In the first study we determined classification agreement among experts without standardized classification criteria. This study involved participant recruitment, video recording and site classification; internationally recognized external rater classification without standardized criteria of patients from masked video recordings; and team rater classification of patients from other centers. In the subsequent study we used a 4-stage Delphi Process

### Key Points

**Question** Are guidelines needed for patient classification for research on adductor spasmodic dysphonia (ADSD), abductor SD (ABSD), voice tremor (VT), and muscular tension dysphonia (MTD)?

**Findings** In this observational study of 178 patients with spasmodic dysphonia and other voice disorders, experts in SD examined uniform video recordings of speech and nasolaryngoscopy examinations of patients with these voice disorders without classification guidelines and had poor interrater agreement that did not relate to clinical classifications. Consequently, a Delphi panel and 46 community experts reached consensus on the Spasmodic Dysphonia Attributes Inventory (SDAI).

**Meaning** The SDAI may be useful for selecting patients with SD, VT, or MTD for research on pathophysiology and treatment outcomes.

to reach consensus on attributes for classifying patients as having ADSD, ABSD, MTD, or VT (eFigure in the [Supplement](#)).

### First Study: Participant Recruitment

Voice care teams who regularly saw patients with SD from Emory University, Washington University in St Louis (WashU), The Medical College of Wisconsin (MCW), and Mayo Clinic Arizona (Mayo-Arizona) recruited participants. Each followed the same protocol and used the same consent form for informed written consent after local institutional review board and National Institute of Neurological Disorders and Stroke–National Institutes of Health approval.

At initial evaluation or on return for reinjection of botulinum toxin at least 2 months after previous injection, patients were approached for participation. Selection criteria were: aged 18 years or older, no known cause for the voice disorder, no medical condition precluding nasolaryngoscopy, no laryngeal structural disease (eg, lesions), no surgery for SD, and diagnosis of ADSD, ABSD, VT, MTD, vocal fold paralysis or paresis (VFP), or other voice disorders (psychogenic dysphonia or hypophonia with Parkinson disease).

### First Study: Clinical Site Classification

After informed written consent, a laryngologist used flexible nasolaryngoscopy to exclude other laryngeal diseases/disorders and diagnose any patient who had vocal fold paralysis or paresis. Nasal topicalization was only used in 2 centers, MCW and Emory. Only participants classified by the laryngologist and SLP as having either ADSD, ABSD, or VT received a medical history and neurological examination. The participant's clinical site classification was entered by the voice team based on their patient knowledge and nasolaryngoscopy by the laryngologist from 11 alternatives: ADSD, ABSD, VT, MTD, ADSD with VT, ABSD with VT, ADSD with MTD, ADSD with VT and MTD, VT with MTD, VFP, or other.

### First Study: Speech and Laryngeal Recordings

Each center used the same equipment and procedures. A head-held microphone was placed at a 45° angle 2 inches off the midline of the mouth to reduce aspiration noise and maintain a

constant amplification level during digital video recording. After the examiner stated the date, patient number, and center, the participant repeated 10 ADSD sentences containing glottal stops in vowels to elicit breaks (eMethods in the Supplement). If the participant missed words in a sentence, the examiner repeated it until the sentence was said correctly while continuing to record. Participants also repeated 10 ABSD sentences containing voiceless consonants (*p, t, k, f, h, s, th*) before vowels, followed by 6 whispered sentences, prolonged vowels, and 2 shouted phrases (eMethods in the Supplement).

The nasolaryngoscope was first placed above the epiglottis to examine supraglottic movement abnormalities by recording posterior tongue and pharyngeal wall movement during repetition of a sniff followed by vowel /i/ repeated 3 times. Two cycles of quiet respiration were recorded. The scope then was moved downward to view the vocal folds during 3 sniff and /i/ repetitions. Next, 5 repetitions of a glottal stop and /i/ were repeated twice, followed by 5 repetitions of the word “see” twice. The participant whistled “Happy Birthday,” and twice prolonged /i/ for 10 seconds. Four sentences were repeated and 2 phrases were shouted (eMethods in the Supplement).

Stroboscopy during a prolonged vowel was not required during nasolaryngoscopy because vibratory tracking was often disrupted by instabilities owing to tremor, spasms, or severe aperiodicity. However, all centers included stroboscopy in the recorded examination whenever possible.

The deidentified digital speech and nasolaryngoscopy recordings were mailed on an encrypted drive to James Madison University, checked for sound, video quality, and task inclusion. The speech and nasolaryngoscopy tasks were extracted into separate video clips for review.

### First Study: Classification of Video Recordings

Two sets of ratings were conducted to assess interrater classification agreement from viewing the speech and nasolaryngoscopy recordings. First, 4 internationally recognized SD experts (C.L.L., C.S., G.B., and M.E.S.) not from the 4 voice centers, served as external raters for some of the first 75 participants. Raters 1 and 2 were SLPs, and raters 3 and 4 were laryngologists. Each reviewed the entire speech and nasolaryngoscopy unclipped video recordings without knowledge of the patient’s diagnosis. They were instructed to select the best classification from the same 11 categories used for site classification (eTable 1 in the Supplement).

Second, 12 team members from the 4 voice centers independently reviewed video recordings of 50 participants seen at other centers. The laryngologists and SLPs reviewed clips of tasks from the speech and nasolaryngoscopy recordings whereas the neurologists only reviewed speech tasks. Each team member independently selected the best classification from the same 11 categories after viewing the speech recording. The laryngologists and SLPs again selected the best classification after viewing nasolaryngoscopy recordings.

### First Study: Statistical Analysis

Interrater agreement was determined using Cohen’s  $\kappa$  coefficient and percent agreement using SAS statistical software (version 9.4, SAS Institute). The distribution of classifications among the ex-

ternal raters was compared using the likelihood ratio  $\chi^2$ . Adequate agreement among the 4 external raters and the site classification, among the 4 external raters, and among pairs of the blinded site team raters was defined as  $\kappa \geq 0.60$  and a percent agreement of 70% or more.

Interrater reliability was determined only for rater pairs who had reviewed 10 or more of the same participants’ video recordings. The agreement using 11 categories was assessed and then collapsed into the following 6 categories: ADSD for participants classified as ADSD regardless of accompanying disorders, ABSD for participants with ABSD regardless of accompanying disorders, VT only for those classified with VT alone, and MTD only for those with MTD alone. Those with both VT and MTD were classified as VT. Classifications of VFP and other (psychogenic or Parkinson disease) were retained.

### Second Study: Delphi Method to Rank Attributes

A 4-stage Delphi method developed consensus on groups of speech and laryngeal attributes for each of 4 main disorders: ADSD, ABSD, VT, and MTD. The aim was to identify groups of attributes considered essential for classification of each disorder rather than defining combinations of disorders. The Delphi Consensus Group (DCG) included 3 team members from each center and an additional otolaryngologist as voting members and the principal investigator and a statistician.

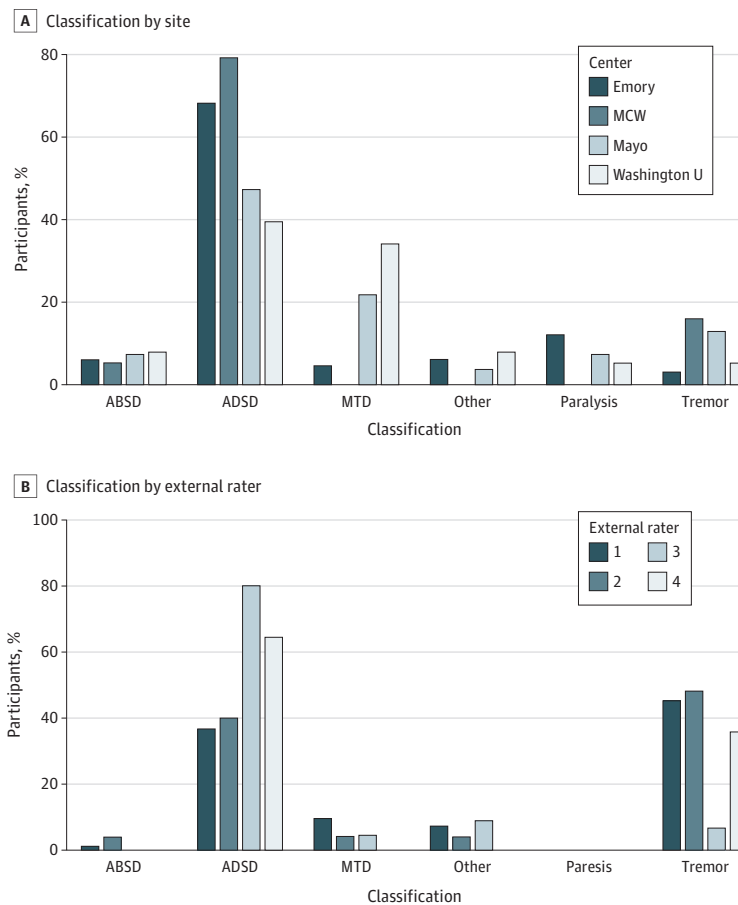
First, each voice team jointly listed an unlimited number of voice and laryngeal attributes for each of the 4 disorders in an Excel file. Second, the DCG convened by conference call and deleted redundancies from the team attribute lists. Next each DCG member independently ranked attributes for each disorder from high to low importance and sent their ranking to the principal investigator. The results were collated and sent to the DCG which convened and eliminated poorly ranked attributes that did not differentiate among disorders.

Third, community SD experts were recruited to independently rank the attributes for each disorder from: the Neurology Study Section of the AAO-HNS, the Voice Disorders Interest Group of the American Speech-Language-Hearing Association and neurologists from the Movement Disorder Society. They were sent an email with a Qualtrics survey link for ranking the attributes for each disorder. They were asked to reorder the attributes based on their perceived order of importance and record their rankings by moving a bar from the top to the bottom of the screen for each disorder. The median rank of the community experts for each attribute per disorder was computed and the percent of experts determined who gave an attribute the same rank as the median. Finally, the DCG reviewed the ranked attributes from the community experts and finalized the attributes for each disorder creating the Spasmodic Dysphonia Attributes Inventory (SDAI).

## Results

Of 197 patients admitted for study, 19 were excluded owing to an unacceptable video recording, lack of a voice disorder, missing sound or tasks, leaving 178 participants; 66 from Emory; 55 from Mayo-Arizona; 38 from WashU; and 19 from MCW. The site clas-

Figure 1. Disorder Classification Distributions by Site and by External Raters



Percentage of participants in categories classified at each center (A) and by external raters (B). The centers are Emory University (Emory), Medical College of Wisconsin (MCW), Mayo Clinic at Scottsdale (Mayo), and Washington University at St Louis (Washington U). The classification categories shown are adductor spasmodic dysphonia (ABSD), abductor spasmodic dysphonia (ASD), muscular tension dysphonia (MTD), other voice disorders such as psychogenic voice disorders (other), vocal fold paralysis or paresis (paralysis), or vocal fold tremor (tremor). The external raters differed in their classification distributions (likelihood ratio  $\chi^2 = 107.66, df = 27, P < .001$ ).

sifications identified 102 (57.2%) with ASD either alone or combined with VT and/or MTD (eTable 1 in the Supplement).

After collapsing the 11 categories into 6, the percentage classified by voice disorder varied across centers (Figure 1A). For example, 12 (20%) or more were diagnosed as MTD at WashU and Mayo-Arizona, none as MTD at MCW, and less than 3 (10%) at Emory. Thus, only 2 of 4 centers identified 20% or more of their participants as having MTD.

**First Study: External Raters and Site Classification**

Percent agreement and Cohen’s  $\kappa$  values between external raters and site classification and was low for 11 and 6 categories (Table 1). For 11 categories, the highest agreement was 34%, with no  $\kappa$  values greater than 0.26. Using 6 categories, the highest percent agreement was 73.3% and the highest  $\kappa$  was 0.40 (Table 1). Thus, none of the external raters had acceptable agreement with the recruiting site classification.

A confusion matrix (eTable 2 in the Supplement) compared the 6 category classifications of the external raters and the sites on 160 patients. Of the 128 patients classified as ASD by sites, only 74 (57.8%) were classified as ASD by external raters, whereas 20 (15.6%) were classified as MTD only and 24 (18.75%) were classified as VT only by raters. Thus, the external raters differentiated between categories more than the sites (Figure 1).

**First Study: Interrater Agreement**

The external raters differed in their distributions of classifications (Figure 1B). Raters 1 and 2 classified 37 and 12 (45% and 48%, respectively) as VT, whereas raters 3 and 4 classified 36 and 9 (80% and 64.3%, respectively) as ASD.

Interrater agreement was low in external rater pairs (Table 1). Using 11 categories, the highest  $\kappa$  was 0.23 and the highest agreement was 38.5% for raters 3 and 4. Most external rater classifications were ASD whereas other categories had zero cases (Figure 1B), likely adversely affecting the  $\kappa$  values. Using 6 categories, the highest agreement was 64% and the highest  $\kappa$  was 0.35 for raters 1 and 3 (Table 1). Although these values increased when using 6 categories, interrater agreement remained unsatisfactory.

Pairs of voice team members classifications using the 6 collapsed categories were examined based on whether they were either: different professionals at the same center; the same profession but different centers; or different centers and professions. Twenty-four rater pairs performed classifications on 10 or more of the same participants’ speech recordings; only 7 rater pairs completed classifications on 10 or more of the same participants after reviewing speech and nasolaryngoscopy recordings. Two team members’ ratings were discarded because they were not conducted independently.

**Table 1. Percent Agreement of Each External Rater With Site Classification and Between External Rater Pairs**

Agreement Pairing	Ratings, No.	Agreement Using 11 Categories, %	$\kappa$ for Agreement Using 11 Categories	Agreement Using 6 Categories, %	$\kappa$ for Agreement Using 6 Categories
<b>Rater agreement with site</b>					
1	76	34.0	0.261	61.3	0.399
2	25	8.0	0.047	12.0	0.047
3	45	28.8	0.108	73.3	0.005
4	14	28.6	0.191	57.1	0.293
<b>Rater pairs</b>					
1 and 2	25	16.0	0.112	36.0	0.234
1 and 3	39	28.0	0.157	64.0	0.348
1 and 4	14	35.0	0.200	43.0	0.104
2 and 3	24	8.0	0.047	17.0	0.068
2 and 4	13	7.7	-0.047	7.6	-0.110
3 and 4	13	38.5	0.235	53.8	-0.013

The  $\kappa$  values among voice team members were also low (<0.60) (Figure 2). Only 2 of 24  $\kappa$  values were above 0.43 for classification based on viewing speech and 3 of 7 based on viewing both speech and nasolaryngoscopy recordings (Figure 2). This indicates that although still unsatisfactory, interrater agreement improved after reviewing both speech and nasolaryngoscopy recordings.

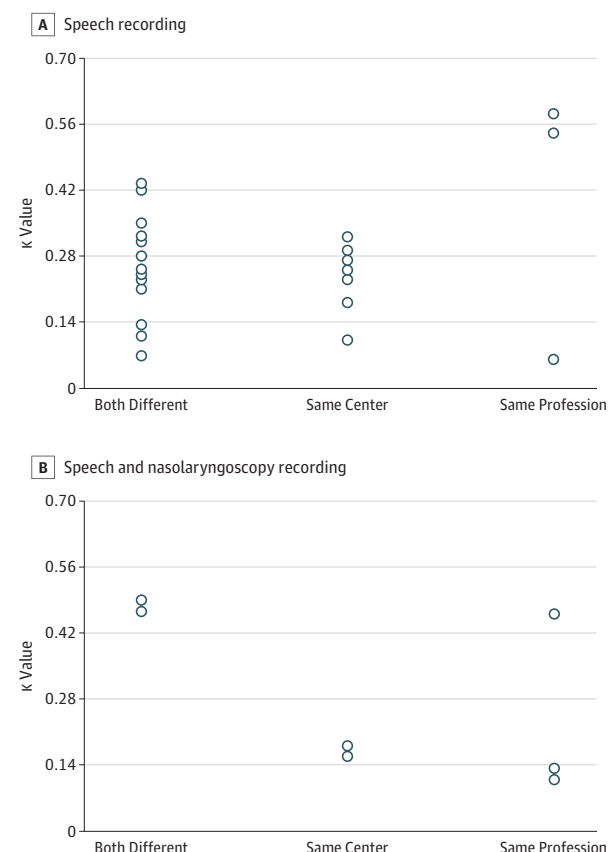
**Second Study: Delphi Method to Identify Attributes**

First, the initial lists of attributes from each center yielded 52 attributes: 16 for ADSD, 15 for ABSD, 11 for MTD, and 10 for VT (Table 2). Second, the DCG reduced these to 42 by deleting 10 attributes that were: not specific to disorders, based on patient reporting, not required for identification, or required electromyography which is unavailable in many voice clinics. The median attribute ranks for each disorder by DCG members are in Table 2. Several nasolaryngoscopy items were retained including AD7 and AD11 to exclude other laryngeal disorders by otolaryngologists by noting, “normal structure and function of the vocal folds” and “intermittent vocal fold/arytenoid hyperadduction” respectively. Items AB2, AB3, MTD2, TR4, TR5, and TR7 were retained for similar reasons.

Third, the 81 community experts contacted included: 35 otolaryngologists, 34 SLPs, and 12 neurologists. Forty-nine (60.5%) responded including 17 otolaryngologists, 26 SLPs, and 6 neurologists. Experts were asked to rate attributes only on disorders familiar to them. Forty-five ranked ADSD attributes, 38 ABSD, 41 MTD, and 33 VT. The median attribute ranks by community experts for each disorder and the percent of community experts who gave an attribute the median ranking are in Table 2. Overall the ordering of attributes by the community experts agreed with the DCG. The highest-ranked items for each disorder had the highest agreement among experts.

Fourth, the DCG reviewed the community experts’ results and concluded that a smaller number of attributes could identify each disorder. They agreed that speech attributes could only identify a probable disorder whereas nasolaryngoscopy items were required to exclude other laryngeal disorders. The

**Figure 2. Interrater Agreement After Viewing Speech Recording and Speech and Nasolaryngoscopy Recordings**



A, Interrater agreement after viewing speech recording. B, Interrater agreement after viewing speech and nasolaryngoscopy recordings. The  $\kappa$  scores are based on 2 independent raters viewing at least 10 of the same participants. The  $\kappa$  scores were provided when both viewers were either from different professions and from different voice centers (both different), from the same voice center but different professions (same center), or from the same profession but from different voice centers (same profession).



**Table 2. Median Rankings of Attributes by the Delphi Consensus Group (DCG) and the Community Experts, and the Percent Agreement Among the Community Experts With the Median Rank**

Original Item Number in DCG List	Attribute	DCG Median Rank	Community Experts Median Rank	Agreement Among Community Experts, %
AD1 <sup>a</sup>	Intermittent glottal stops (vowel breaks) in vowels on all voiced sentences	1	1	53
AD2 <sup>a</sup>	Strain-strangled, effortful, tight voice quality	2	2	44
AD14 <sup>a</sup>	Patient report of speaking effort	6	3	47
AD3 <sup>a</sup>	Symptoms reduced during whisper	7	6	27
AD10	Symptoms reduced with voiceless consonants	5	5	33
AD4	Symptoms reduced during singing	7	6	33
AD7 <sup>a</sup>	Normal structure and symmetry of vocal folds at rest	9	7	18
AD8	Patient reports symptoms worse on telephone	7	8	27
AD5	Symptoms reduced during shouting or yelling	9	9	31
AD6	Symptoms reduced during laughing or crying	8	9	36
AD12	Symptoms reduced during high pitch phonation	7	11	40
AD11 <sup>a</sup>	Intermittent vocal fold/arytenoid hyper adduction	12	11	49
AB1 <sup>a</sup>	Intermittent breathy breaks on voiceless consonants before vowels in speech	1	1	97
AB7 <sup>a</sup>	Symptoms most evident during connected speech	4	2	76
AB10 <sup>a</sup>	Few symptoms on prolonged vowels	4	3	55
AB2 <sup>a</sup>	Intermittent abductor spasm of vocal folds/arytenoids during speech	8	4	50
AB9	Patient reports of "running out of air" while speaking	6	5	39
AB8	Patient report of symptoms worse on the telephone	6	7	29
AB3 <sup>a</sup>	Normal structure and symmetry of vocal folds at rest	7	7	29
AB4	Symptoms reduced during whisper	8	8	26
AB5	Symptoms reduced during singing	7	9	32
AB14	Patient reports avoidance of specific words	8	8	32
AB6	Symptoms may reduce during shouting or yelling	11	11	58
MTD1 <sup>a</sup>	Continuous effortful strained voice quality	1	1	78
MTD3 <sup>a</sup>	Similar voice quality abnormalities on all types of sounds	3	2	68
MTD2 <sup>a</sup>	Continuous supraglottic compression obscuring vocal fold view during voice production	2	3	44
MTD4 <sup>a</sup>	Absence of phonatory breaks	4	5	32
MTD6	Patient complaint of pain in the throat area that may increase with laryngeal palpation	6	5	34
MTD5	Voice quality does not improve during singing or shouting	7	6	32
MTD8	Symptoms may improve with instruction or laryngeal manipulation	8	6	32
MTD9	Signs of increased neck tension and elevated laryngeal position during speech	8	7	27
MTD10	Patient may report rapid onset of disorder	9	9	37
MTD12	Symptoms may not improve during whisper	11	10	37
MTD11	Patient may report that symptoms go away for more than a day	10	11	51
TR1 <sup>a</sup>	Regular pitch and/or amplitude oscillation during a sustained vowel	1	1	97
TR4 <sup>a</sup>	Tremor on multiple sites on nasolaryngoscopy (eg, vocal folds, pharynx, tongue)	2	2	88
TR3	Not affected by phonetic content	3	3	52
TR2	Voice tremor (shaky voice) less obvious during connected speech	3	4	52
TR5 <sup>a</sup>	Laryngeal position may show regular bobbing	5.5	5	61
TR8	Severe cases may have regular glottal stops during vowels	6	6	61
TR6	Patient may have head and/or hand tremor	6	7	55
TR7	May have visible tremor at rest in affected sites	6.5	8	76

Abbreviations: AB, abductor spasmodic dysphonia; AD, adductor spasmodic dysphonia; MDT, muscular tension dysphonia; TR, vocal fold tremor.

<sup>a</sup> Items were included by DCG at stage 4.

Figure 3. The Spasmodic Dysphonia Attribute Inventory Form for Identifying Disorder Classification

**Spasmodic Dysphonia Attribute Inventory Score Sheet (v. 1.6)**

Please check whether an attribute is present (✓) or not (–) for the patient being scored. The designation of the patient’s predominant symptom category will be based on how many of the attributes are checked for each category.

**Spasmodic Dysphonia Adductor Type**

Item	Description	Check if present
1	Intermittent glottal stops (vowel breaks) in vowels on voiced sentences	
2	Strain-strangled, effortful, tight voice quality	
3	Patient report of speaking effort	
4	Symptoms reduced during whisper	
5	Normal structure and symmetry of the vocal folds at rest	
6	Intermittent vocal fold or arytenoid hyperadduction	

**Spasmodic Dysphonia Abductor Type**

7	Intermittent breathy breaks in voiceless consonants before vowels	
8	Symptoms most evident during connected speech	
9	Few symptoms on prolonged vowels	
10	Normal structure and symmetry of the vocal folds at rest	
11	Intermittent abductor spasm of the vocal folds or arytenoids during speech	

**Muscular Tension Dysphonia**

12	Continuous effortful strained voice quality	
13	Similar voice quality abnormalities on all types of sounds	
14	Absence of phonatory breaks	
15	Continuous supraglottic compression obscuring the vocal folds during voice production	

**Vocal Tremor**

16	Regular pitch and/or amplitude oscillation during a sustained vowel	
17	Tremor seen by nasolaryngoscopy that may involve several sites (vocal folds, pharynx, tongue)	
18	Laryngeal position may show bobbing during voice	

**OTHER DIAGNOSIS**

VF paresis or paralysis	Check this box ONLY if vocal fold paresis or paralysis is present	
Other	Check this box ONLY if NO other diagnosis is checked (eg, psychogenic)	

Users check those items observed during speech and nasolaryngoscopy examinations. Those disorders with half or more of the items checked should be considered for classification.

DCG identified 18 attributes for the SDAI: 11 speech attributes and 7 nasolaryngoscopy items across the 4 disorders (Figure 3).

## Discussion

In the first study, experienced specialists without standardized criteria had poor agreement classifying patients with SD and related voice disorders. Thus, research studies reporting patients with either ADSD, ABSD, VT, or MTD may have groups with dissimilar characteristics. Further, the distribution of patients differed across centers (Figure 1A), which may reflect both referral and classification differences among centers. For example, some centers had 30% patients with MTD whereas others had none or very few with MTD (Figure 1A). Of 102 patients with some form of ADSD, only 35 (19.7%) had ADSD

alone; 27 (15.1%) also had VT, 31 (17.4%) also had MTD, and 9 (5.0%) had all 3 disorders (eTable 1 in the Supplement). Thus, patients with ADSD varied in their characteristics. Also, external raters tended to classify patients differently from the recruiting sites; the external raters classified 74 of 160 (46.25%) as ADSD whereas the sites classified 128 of 160 (80%) as ADSD (eTable 2 in the Supplement). This suggests that some experts tend to group all patients with ADSD as similar whereas others differentiate tremor and MTD from ADSD.

The second study used the Delphi process involving the DCG panel and community experts to establish consensus for identifying and ranking attributes characteristic of these different disorders. Agreement was high on the top attributes for ABSD, 97% for “intermittent breathy breaks” AB1 and 76% for “symptoms most evident in speech” AB7 and for VT, 97% for “pitch and amplitude modulation on prolonged vowels” TR1, and 88% on “tremor on multiple sites

on nasolaryngoscopy” TR4 (Table 2). However, for MTD, the top attributes had lower agreement at 78% on “continuous effortful strained voice” MTD1 and 68% on “similar voice abnormalities on all types of sounds” MTD3. Agreement was lowest for ADSD at 53% for “intermittent glottal stops” AD1 and 47% for “patient report of speaking effort” AD14. Thus, the characteristics of ABSD and VT were better delineated by experts than those for ADSD and MTD.

The poor classification agreement among experts on patient characteristics found in the first study is problematic for research. The lack of agreement among experts on ADSD characteristics can impair clinical trials owing to participant cohort heterogeneity. Evaluation of whether using the attributes from the SDAI to classify patients improves the agreement of patient classification across experts and centers is the next step needed for future research on SD and related disorders.

### Limitations

This study had limitations. Most important was that the raters used video recordings of speech and nasolaryngoscopic examinations including stroboscopy whenever feasible but did not have extensive medical background on the patients. Therefore, this study did not assess the agreement between centers for clinical diagnosis of ADSD, ABSD, VT, and MTD. Rather we only examined agreement when raters were asked to classify patients for research. Also, the initial site

classification for many patients with ADSD was prior to botulinum toxin treatment, which alters clinical manifestations. Further, some participants may have had injections as recently as 2 months earlier, which may not be enough time for full manifestations to recur. We excluded participants that had no symptom manifestations at time of recruitment.

## Conclusions

Our findings demonstrate that classification by experts without benefit of standardized criteria failed to be consistent across independent experts and different recruiting sites. There was poor agreement among independent experts based on viewing extensive speech and nasolaryngoscopy recordings. It was unexpected that members of the same voice care team had unacceptable interrater agreement when conducting ratings independently both for speech and nasolaryngoscopy. Further, agreement was no better when raters were members of different voice care teams but from the same specialty. The 4-stage Delphi method permitted experts to reach consensus on attributes for identifying types of SD and related disorders. This checklist of attributes may allow investigators to develop homogenous patient cohorts for multicenter studies of ADSD, ABSD, VT, or MTD.

### ARTICLE INFORMATION

**Accepted for Publication:** March 20, 2018.

**Published Online:** June 21, 2018.  
doi:10.1001/jamaoto.2018.0644

**Author Affiliations:** Department of Communication Sciences and Disorders, James Madison University, Harrisonburg, Virginia (Ludlow, Barone, Drulia); Department of Mathematics and Statistics, James Madison University, Harrisonburg, Virginia (Domangue, Sharma); Department of Neurology, Emory University School of Medicine, Atlanta, Georgia (Jinnah); Neurology, Radiology, Neuroscience, Physical Therapy and Occupational Therapy, Washington University School of Medicine, St Louis, Missouri (Perlmutter); Head and Neck Surgery, Ronald Reagan University of California Los Angeles Medical Center, Los Angeles (Berke); Communication Sciences and Disorders, Jacksonville University, Jacksonville, Florida (Sapienza); Division of Otolaryngology-Head and Neck Surgery, University of Utah School of Medicine, Salt Lake City (Smith); Division of Laryngology and Professional Voice, Medical College of Wisconsin, Milwaukee (Blumin, Kalata); Department of Neurology, Medical College of Wisconsin, Milwaukee (Blindauer); Emory Voice Center, Department of Otolaryngology-Head and Neck Surgery, Emory University, Emory University, Atlanta, Georgia (Johns, Hapner); Department of Otolaryngology-Head and Neck Surgery, Washington University School of Medicine, St Louis, Missouri (Harmon, Paniello); Department of Neurology, Mayo Clinic College of Medicine, Mayo Clinic Arizona, Scottsdale, Arizona (Adler); Department of Otolaryngology, Mayo Clinic College of Medicine, Mayo Clinic Arizona, Scottsdale,

Arizona (Crujido, Lott, Bansberg); Department of Neurological Sciences, Rush Medical College, Chicago, Illinois (Stebbins); Division of Laryngology, Keck Medicine of University of Southern California, Los Angeles (Johns, Hapner); Communication Sciences and Disorders, Curry School of Education, University of Virginia, Charlottesville (Barone); Davies School of Communication Sciences and Disorders, Harris College of Nursing and Health Sciences, Texas Christian University, Fort Worth (Drulia).

**Author Contributions:** Dr Ludlow had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Study concept and design:** Ludlow, Domangue, Jinnah, Perlmutter, Sapienza, Blumin, Johns, Adler, Lott, Stebbins.

**Acquisition, analysis, or interpretation of data:** Ludlow, Domangue, Sharma, Jinnah, Perlmutter, Berke, Smith, Blumin, Kalata, Blindauer, Johns, Hapner, Harmon, Paniello, Adler, Crujido, Bansberg, Barone, Drulia, Stebbins.

**Drafting of the manuscript:** Ludlow, Domangue, Sharma, Sapienza, Barone, Stebbins.

**Critical revision of the manuscript for important intellectual content:** Ludlow, Sharma, Jinnah, Perlmutter, Berke, Smith, Blumin, Kalata, Blindauer, Johns, Hapner, Harmon, Paniello, Adler, Crujido, Lott, Bansberg, Barone, Drulia, Stebbins.

**Statistical analysis:** Ludlow, Domangue, Sharma, Stebbins.

**Obtained funding:** Ludlow, Jinnah, Blumin.

**Administrative, technical, or material support:** Ludlow, Jinnah, Perlmutter, Berke, Sapienza, Blumin, Johns, Crujido, Lott, Bansberg, Barone.

**Study supervision:** Ludlow, Jinnah, Lott.

**Conflict of Interest Disclosures:** All authors have completed and submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest and none were reported.

**Funding/Support:** This study was supported by the National Spasmodic Dysphonia Association and a grant to the Dystonia Coalition from the Office of Rare Diseases Research in the National Institutes of Health (NIH) National Institute of Translational Sciences (grant number U54 TRO01456) and the National Institute of Neurological Disorders and Stroke (grant number U45 NS065701) from the NIH.

**Role of the Funder/Sponsor:** the National Spasmodic Dysphonia Association and the National Institutes of Health had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

### REFERENCES

- Nutt JG, Muentner MD, Aronson A, Kurland LT, Melton LJ III. Epidemiology of focal and generalized dystonia in Rochester, Minnesota. *Mov Disord*. 1988;3(3):188-194.
- Blitzer A, Brin MF, Fahn S, Lovelace RE. Clinical and laboratory characteristics of focal laryngeal dystonia: study of 110 cases. *Laryngoscope*. 1988;98(6 Pt 1):636-640.
- Sapienza CM, Murry T, Brown WS Jr. Variations in adductor spasmodic dysphonia: acoustic evidence. *J Voice*. 1998;12(2):214-222.



4. Sapienza CM, Walton S, Murry T. Acoustic variations in adductor spasmodic dysphonia as a function of speech task. *J Speech Lang Hear Res.* 1999;42(1):127-140.
5. Edgar JD, Sapienza CM, Bidus K, Ludlow CL. Acoustic measures of symptoms in abductor spasmodic dysphonia. *J Voice.* 2001;15(3):362-372.
6. Whurr R, Lorch M. Review of differential diagnosis and management of spasmodic dysphonia. *Curr Opin Otolaryngol Head Neck Surg.* 2016;24(3):203-207.
7. Sulica L, Louis ED. Clinical characteristics of essential voice tremor: a study of 34 cases. *Laryngoscope.* 2010;120(3):516-528.
8. Morrison MD, Rammage LA, Belisle GM, Pullan CB, Nichol H. Muscular tension dysphonia. *J Otolaryngol.* 1983;12(5):302-306.
9. Creighton FX, Hapner E, Klein A, Rosen A, Jinnah HA, Johns MM. Diagnostic delays in spasmodic dysphonia: a call for clinician education. *J Voice.* 2015;29(5):592-594.
10. Roy N. Differential diagnosis of muscle tension dysphonia and spasmodic dysphonia. *Curr Opin Otolaryngol Head Neck Surg.* 2010;18(3):165-170.
11. Ludlow CL, Adler CH, Berke GS, et al. Research priorities in spasmodic dysphonia. *Otolaryngol Head Neck Surg.* 2008;139(4):495-505.
12. Johnson DM, Hapner ER, Klein AM, Pethan M, Johns MM III. Validation of a telephone screening tool for spasmodic dysphonia and vocal fold tremor. *J Voice.* 2014;28(6):711-715.
13. Daraei P, Villari CR, Rubin AD, et al. The role of laryngoscopy in the diagnosis of spasmodic dysphonia. *JAMA Otolaryngol Head Neck Surg.* 2014;140(3):228-232.
14. Roy N, Whitchurch M, Merrill RM, Houtz D, Smith ME. Differential diagnosis of adductor spasmodic dysphonia and muscle tension dysphonia using phonatory break analysis. *Laryngoscope.* 2008;118(12):2245-2253.
15. Houtz DR, Roy N, Merrill RM, Smith ME. Differential diagnosis of muscle tension dysphonia and adductor spasmodic dysphonia using spectral moments of the long-term average spectrum. *Laryngoscope.* 2010;120(4):749-757.
16. Parnes SM, Lavorato AS, Myers EN. Study of spastic dysphonia using videofiberoptic laryngoscopy. *Ann Otol Rhinol Laryngol.* 1978;87(3 Pt 1):322-326.
17. Belisle GM, Morrison MD. Anatomic correlation for muscle tension dysphonia. *J Otolaryngol.* 1983;12(5):319-321.
18. Blitzer A, Brin MF, Stewart CF. Botulinum toxin management of spasmodic dysphonia (laryngeal dystonia): a 12-year experience in more than 900 patients. *Laryngoscope.* 1998;108(10):1435-1441.
19. Ludlow CL, Naunton RF, Sedory SE, Schulz GM, Hallett M. Effects of botulinum toxin injections on speech in adductor spasmodic dysphonia. *Neurology.* 1988;38(8):1220-1225.
20. Berke GS, Blackwell KE, Gerratt BR, Verneil A, Jackson KS, Sercarz JA. Selective laryngeal adductor denervation-reinnervation: a new surgical treatment for adductor spasmodic dysphonia. *Ann Otol Rhinol Laryngol.* 1999;108(3):227-231.
21. Chhetri DK, Mendelsohn AH, Blumin JH, Berke GS. Long-term follow-up results of selective laryngeal adductor denervation-reinnervation surgery for adductor spasmodic dysphonia. *Laryngoscope.* 2006;116(4):635-642.
22. Hertegård S, Granqvist S, Lindestad PA. Botulinum toxin injections for essential voice tremor. *Ann Otol Rhinol Laryngol.* 2000;109(2):204-209.
23. Roy N, Bless DM, Heisey D, Ford CN. Manual circumlaryngeal therapy for functional dysphonia: an evaluation of short- and long-term treatment outcomes. *J Voice.* 1997;11(3):321-331.
24. Roy N, Ford CN, Bless DM. Muscle tension dysphonia and spasmodic dysphonia: the role of manual laryngeal tension reduction in diagnosis and management. *Ann Otol Rhinol Laryngol.* 1996;105(11):851-856.
25. Chhetri DK, Merati AL, Blumin JH, Sulica L, Damrose EJ, Tsai VW. Reliability of the perceptual evaluation of adductor spasmodic dysphonia. *Ann Otol Rhinol Laryngol.* 2008;117(3):159-165.
26. Nash EA, Ludlow CL. Laryngeal muscle activity during speech breaks in adductor spasmodic dysphonia. *Laryngoscope.* 1996;106(4):484-489.
27. Cannito MP, Chorna LB, Kahane JC, Dworkin JP. Influence of consonant voicing characteristics on sentence production in abductor versus adductor spasmodic dysphonia. *J Voice.* 2014;28(3):394e313-322.
28. Patel RR, Liu L, Galatsanos N, Bless DM. Differential vibratory characteristics of adductor spasmodic dysphonia and muscle tension dysphonia on high-speed digital imaging. *Ann Otol Rhinol Laryngol.* 2011;120(1):21-32.
29. Okoli C, Pawlowski SD. The Delphi method as a research tool: an example, design considerations and applications. *Inf Manage.* 2004;42:15-29.
30. Hsu C-C, Sandford BA. The Delphi Technique: Making sense of Consensus. *Pract Assess Res Eval.* 2007;12(10):1-8.
31. Skulmoski GJ, Hartman FT, Kahn J. The Delphi method for graduate research. *J Inf Technol Educ.* 2007;6:1-21.
32. de Ru MH, Teunissen QG, van der Lee JH, et al. Capturing phenotypic heterogeneity in MPS I: results of an international consensus procedure. *Orphanet J Rare Dis.* 2012;7:22.
33. Wildfire JJ, Gergen PJ, Sorkness CA, et al. Development and validation of the Composite Asthma Severity Index--an outcome measure for use in children and adolescents. *J Allergy Clin Immunol.* 2012;129(3):694-701.
34. Mitha A, Boulyana M, Hue V, Pruvost I, Martinot A, Dubos F; European French-speaking expert group. Consensus in diagnostic definitions for bone or joint infections in children by a Delphi method with European French-speaking experts. *Acta Paediatr.* 2012;101(8):e350-e356.
35. Mamisch N, Brumann M, Hodler J, Held U, Brunner F, Steurer J; Lumbar Spinal Stenosis Outcome Study Working Group Zurich. Radiologic criteria for the diagnosis of spinal stenosis: results of a Delphi survey. *Radiology.* 2012;264(1):174-179.
36. Bernard TJ, Manco-Johnson MJ, Lo W, et al. Towards a consensus-based classification of childhood arterial ischemic stroke. *Stroke.* 2012;43(2):371-377.
37. Gabel MJ, Foster NL, Heidebrink JL, et al. Validation of consensus panel diagnosis in dementia. *Arch Neurol.* 2010;67(12):1506-1512.