Validation of the Dyspnea Index in Adolescents With Exercise-Induced Paradoxical Vocal Fold Motion

Vanessa De Guzman, MS; Catherine L. Ballif, MA, CCC-SLP; Rie Maurer, MA; Christopher J. Hartnick, MD, MSEpi; Nikhila Raol, MD

**IMPORTANCE** Paradoxical vocal fold motion (PVFM) affects almost 1 million adolescents in the United States. However, to date, no disease-specific objective measure exists to assess symptom severity and response to treatment in adolescents with exercise-induced PVFM.

**OBJECTIVES** To validate the Dyspnea Index (DI) quality-of-life instrument (previously validated for adults with breathing disorders) in children aged 12 to 18 years with exercise-induced PVFM and to determine the minimum significant DI change corresponding to patient-reported or caregiver-reported improvement or worsening of symptoms.

**DESIGN, SETTING, AND PARTICIPANTS** A longitudinal study of 56 patients (age range, 12-18 years) diagnosed as having exercise-induced PVFM and their caregivers from February 1, 2013, to September 30, 2013, in an outpatient pediatric otolaryngology office practice.

**INTERVENTIONS** The DI was administered to patients and caregivers, with items modified to reflect the perspective of caregivers.

**MAIN OUTCOMES AND MEASURES** Appropriate DI change was measured to reflect improvement or worsening of symptoms. Test-retest reliability was accomplished by having a subset of patients and caregivers complete the instrument twice within 2 weeks before therapy. Internal consistency was assessed by calculation of Cronbach α. Discriminant validity and convergent validity were determined by comparing DIs with assessment of global change in symptoms.

**RESULTS** The patient and caregiver mean (SD) DI changes were −12.9 (9.6) and −14.7 (9.3), respectively ($P < .001$ for both). Reliability was established by test-retest analysis with an intraclass correlation coefficient of 0.8 and by calculation of Cronbach α = 0.80, demonstrating internal consistency. Discriminant validity was determined by assessing for a significant DI change when patients globally perceived that a change existed after treatment ($P < .001$). A DI change of 8 or higher ($P < .001$) correlated with patient-reported significant change. Convergent validity was demonstrated by evaluating for significant DI change when no change was reported following treatment ($P < .001$).

**CONCLUSIONS AND RELEVANCE** The DI appears to be a valid and reliable instrument to assess quality of life in exercise-induced pediatric PVFM. A DI change of 8 or higher seems significant. This instrument can serve as an objective tool to assess change in exercise-induced pediatric PVFM following speech therapy.
paradoxical vocal fold motion (PVFM) is defined as inappropriate adduction of the vocal folds with inspiration. Initially described as Munchausen stridor in 1974 by Patterson et al,1 PVFM has been referred to by various other names, including vocal cord dysfunction, episodic laryngospasm, episodic laryngeal dyskinesia, psychosomatic stridor, emotional laryngeal asthma, and irritable larynx.2-3 Improper laryngeal closure may result in partial to severe airway obstruction with resulting cough, shortness of breath, and a perception of throat tightness.

Typical PVFM episodes occur with exercise or with recurrent stressful activity. It may manifest with various symptoms: a meta-analysis4 of 288 articles on PVFM studying 1020 patients demonstrated prevalent symptoms of dyspnea (73%), wheezing (36%), perceivable stridor (28%), cough (25%), chest tightness (25%), throat tightness (22%), and voice change (12%). These symptoms typically abate when the inciting activity is stopped but recur once the activity is resumed. While the diagnosis can be confirmed with flexible fiber-optic laryngoscopy or spirometry, the results are typically normal when the patient is asymptomatic and not performing the activity that causes onset. When symptoms are present at rest, an underlying psychiatric disorder must be considered. Maturo et al5 demonstrated that 30% of patients with PVFM were found to have a psychiatric component. If the patient is symptomatic or symptoms can be reproduced, laryngoscopy may reveal adduction of the anterior two-thirds of the true vocal folds, with a diamond-shaped posterior glottis chink with inspiration,6 and spirometry may show variable extrathoracic obstruction on inspiration, although this is not particularly sensitive or specific.4 Therefore, a detailed clinical history is most often used to make the diagnosis. Several causes have been proposed for this disorder, and numerous classification schemes have been developed in an attempt to group by etiology (eg, psychiatric, reflux, and neurologic origins, among other less common causes).

It is estimated by the National Heart, Lung, and Blood Institute that almost 40% of the 22 million patients, or 8.8 million people, diagnosed as having asthma in fact have PVFM.4 This condition is often misdiagnosed as asthma, with studies7-8 showing that 10% to 40% of patients diagnosed as having severe refractory asthma actually have PVFM. In the pediatric population, it is estimated that more than 5% of children aged 10 to 18 years in the United States experience PVFM.9 Given US Census Bureau statistics that document 18 million children aged 10 to 18 years in 2010, as many as 916,686 children aged 12 to 18 years may have PVFM.10 Among children with exercise-induced dyspnea, Abu-Hasan et al11 showed a 10% prevalence of PVFM, and studies12,13 of high-level athletes have reported a prevalence of 3% to 5%.

Different treatment modalities have been used, with varying success. Antireflux medications have been prescribed, but limited evidence exists to suggest significant benefit.2 Psychiatric medications may be useful for underlying Diagnostic and Statistical Manual of Mental Disorders. Surgery has a limited role, with botulinum toxin injection used for refractory cases.7 Speech therapy remains the mainstay of treatment, with greater than 80% success noted.14-16 However, to date, no validated, disease-specific, objective quality-of-life measure in the pediatric population exists to document significant changes before and after any treatment modality. The development of a method by which to assess symptom severity and response to treatment can lead to best-practice algorithms via randomized clinical trials, thereby avoiding inappropriate treatment and unnecessary pharmacotherapy in these patients.

The Dyspnea Index (DI), an instrument recently developed by Gartner-Schmidt et al,17 was designed and validated to provide clinicians and patients with a reliable and valid clinical assessment tool to evaluate overall symptom severity of dyspnea in patients with upper airway pathologies. It is a patient-centered, psychometrically robust questionnaire that can be used for assessment of initial severity of dyspnea symptoms related to the upper airway and can be used as a follow-up assessment tool to evaluate treatment outcomes and drive evidence-based medical decisions. It is a 10-item validated instrument scored from 0 to 40, where low totals reflect fewer dyspnea-related symptoms and higher totals reflect more dyspnea-related symptoms (or a worse dyspnea-specific or disease-specific quality of life). Gartner-Schmidt et al evaluated 317 adult patients for various aspects of psychometric assessment of this instrument and found it to be internally consistent (Cronbach α = 0.91), with excellent test-retest reliability (r = 0.79) and significant discriminant validity (P < .001 by Mann-Whitney test). Because of the design of the DI as a brief questionnaire, we believed that this instrument was ideal for use in the exercise-induced pediatric population with PVFM. The objective of this study was to validate the DI in children aged 12 to 18 years with exercise-induced PVFM seen at a pediatric airway center. The secondary objective was to determine the minimum significant DI change corresponding to patient-reported or caregiver-reported improvement or worsening of symptoms. The DI used in this study is a 10-item, disease-specific quality-of-life instrument that was modified for children from the previously validated instrument for adults with dyspnea.17 Because of the age of patients, this survey was given in its native form for self-administration in adolescents with PVFM. In addition, the instrument was modified for administration to caregivers as the child’s proxy. The totals range from 0 to 40, with each item scored on a 4-point Likert-type scale. Higher totals indicate a poorer disease-specific quality of life, while lower totals indicate a better disease-specific quality of life.

Methods

Study Design

Approval was obtained for this study from the Massachusetts Eye and Ear Infirmary Institutional Review Board. Verbal consent was obtained from all caregivers and patients. Eighty-one adolescent patients aged 12 to 18 years with symptoms of PVFM were contacted from February 1, 2013, to September 30, 2013, at the Voice and Speech Laboratory at the Massachusetts Eye and Ear Infirmary. No exclusion criteria were applied based on race/ethnicity or sex. A larger sample size was obtained compared with previous investigations because of the nature of our institution as a tertiary referral center for voice
and breathing disorders. In addition, we presume that increased awareness of the entity of PVFM led to this diagnosis in 81 patients.

A thorough history and physical examination were performed in all patients, including flexible fiber-optic laryngoscopy at rest. A diagnosis was made based on patient and caregiver report of typical PVFM symptoms, including difficulty breathing, throat tightness, stridor, reduced response to rescue inhalers, and blood oxygen saturation levels during an episode, when available. Patients with other causes of dyspnea, including pulmonary, cardiac, and neurologic origins, were excluded from the study. Following the diagnosis, all children were referred to a speech-language pathologist for voice therapy. The DI was administered to patients and caregivers for pretherapy and posttherapy assessment. Fifty-six patients and caregivers completed the study. To analyze test-retest reliability, the last 21 patients completed the DI over the phone or by e-mail before the first appointment with the speech-language pathologist. All answered questionnaires were deidentified and entered into a clinical database. The database was then linked to a software package (SAS 9.3; SAS Institute Inc). Reliability analysis was performed by calculation of Cronbach α to determine internal consistency and by establishment of test-retest reliability via the Bland-Altman method. Discriminant validity was evaluated by assessing for a significant DI change when patients globally perceived that a change existed after treatment. Convergent validity was demonstrated by evaluating for significant DI change when no change was reported following treatment. Cross-correlations between the patient and caregiver DIs were also performed. The minimum DI change correlating with patient-reported and caregiver-reported improvement in symptoms was determined using receiver operating characteristic curve analysis.

Voice Therapy
Voice therapy typically consisted of 1 to 6 sessions of 30 to 60 minutes in duration, occurring every 2 to 3 weeks or as needed. Every attempt was made for the treatment period to overlap with the participant’s athletic season to better facilitate application and successful carryover. Treatment focused on the 3 areas of education, treatment, and carryover. Education typically began during the evaluation visit with explanations of normal laryngeal function, the nature and presentation of PVFM in general and specific to the participant, and the identification of strategies to reduce or control possible contributing factors as well as possible contributing environmental, physical, and emotional factors, followed by initial training in the rescue breathing technique. Also addressed were vocal health and hygiene, including hydration recommendations, behavior modification of phonotraumatic behaviors, and standard dietary precautions and lifestyle modifications for laryngopharyngeal reflux, depending on time and need. The first therapy session focused on training and application in the 2 areas of laryngeal control exercises and laryngeal massage techniques.

Progression through the voice therapy protocol occurs via 3 levels. Time in each phase will vary between individuals. Level 1 is achieved when patients are able to successfully and confidently apply laryngeal control exercises, including res-
Reliability

The reliability of the instrument was assessed by calculating Cronbach’s α, measuring the internal consistency among test items. This was determined to be 0.80 (α > 0.55 was deemed acceptable). Test-retest reliability was demonstrated by administering the DI to a subpopulation of 21 patients and caregivers twice within 2 weeks before therapy. The intraclass correlation coefficient was calculated to be 0.8, which implies substantial reliability.19

Validity

The validity of the survey was evaluated by assessing convergent and discriminant validity. Discriminant validity was demonstrated by the ability to show significant differences among patients for whom a preconceived hypothesis indicated that a difference existed. For all patients who completed the DI, the question was posed as to whether global improvement in their condition had occurred following therapy. For 43 patients who perceived improvement, the mean (SD) decrease in the DI was 15.8 (7.3), demonstrating a statistically significant change (P < .001). Convergent validity was demonstrated by the ability to show no significant difference in pretherapy and posttherapy DIs when no change was reported following treatment. In these 14 patients, a mean (SD) decrease in the DI of 3.6 (3.0) was noted, indicative of convergent validity (P < .001).

Although validity was assessed only for the patients, a sole disagreement was found between patients and caregivers regarding global improvement following treatment. Therefore, the caregiver instrument does not seem to need separate revalidation at this juncture.

The receiver operating characteristic curve for the DI is shown in Figure 2. The area under the curve, which illustrates how well the measurement characterizes persons as having global improvement vs no improvement, is 0.884 (95% CI, 0.771-0.954). Based on this curve, a DI change of 8 indicates the cutoff point at which significant global improvement in symptoms is noted, with a sensitivity of 88.4% and a specificity of 84.6%.

Discussion

The DI is a reliable and validated clinical assessment tool that was recently developed to evaluate adults with respiratory difficulty due to upper airway obstruction. It is ideal for assessing symptom severity and allows for evaluation of efficacy treatment based on worsening or improvement of symptoms. Because of the nature of the instrument, we believed that it could be validated for this discrete disease entity, as well as be translatable to and ideal for use in the adolescent population.

When assessing disease-specific quality-of-life outcomes, it is important to use a valid and reliable instrument because evidence has shown that favorable results, which support the hypothesis of the study, are more likely to be reported with unpublished scales than with validated ones. Particularly in the setting of psychometric analysis, in which subjective responses are used to assess outcomes, a validated tool is necessary to standardize responses and remove bias. As stated before, the DI has been validated in a discrete population with a specific pathology. In applying this instrument to another population, it cannot be assumed that the same psychometric properties hold true for another cohort. Estimates of validity are contingent on the population being tested, as well as the circumstances under which this population is tested. Adolescents presumably place emphasis on different aspects of quality of life compared with adults. Furthermore, it is unknown whether the DI, which was validated for general upper airway obstruction, will be as relevant to a specific disease process, such as PVFM. Therefore, it is necessary to reevaluate this tool in this specific age group with this specific diagnosis and establish validity and reliability. In addition, when translating the instrument into other languages, it is necessary to establish conceptual equivalence (ie, whether different cultures view the problem in the same way). Once this balance is established, item equivalence (which determines whether the specific items are relevant and acceptable in the new population) and semantic equivalence (which refers to the meaning attached to each item) must be addressed. Only then can validity be established in the new language.

When validating an instrument that has no other validated measure with which to compare, construct validity must be assessed by developing a hypothesis regarding the relationship of the instrument scores and the effect of the disease on the target population. With this data set, we predicted that patients with PVFM would demonstrate higher DIs before treatment and that those who noted global improvement in symptoms following treatment would demonstrate lower DIs.

Our results support that the DI accurately assesses symptom severity in pediatric PVFM, with pretreatment DIs being higher for both the patient and the caregiver, indicating more severe symptoms. In addition, posttreatment DIs fell, reflecting improvement in symptoms for almost all patients both by self-assessment and caregiver proxy. Pretreatment and post-
treatment measures showed a statistically significant decrease in DIs after voice therapy ($P < .001$), indicating considerable symptomatic improvement. Furthermore, a DI change of 8 or higher ($P < .001$) after therapy correlated with global improvement in symptoms with greatest sensitivity and specificity. At this cutoff point, the probability of a patient who reported global improvement having a lower DI than a patient who did not report global improvement is about 88%.

The DI is particularly useful in the adolescent population because patients in this age group, as opposed to younger patients, possess the maturity, cognitive skills, and awareness of their condition to be able to answer the questions. While quality-of-life assessment in younger children may be more effective through a caregiver proxy, this is not typically the case once a child reaches adolescence years. Research in pediatric psychology has demonstrated that, at approximately age 12 years, children begin to understand the complexity of illnesses and are able to describe their ailment in less concrete and more abstract terms. In addition, given the effect that the symptoms (eg, shortness of breath with exertion, stridor, and anxiety) have on daily functioning, the adolescent may be the first to notice the issue and seek help. While the caregiver responses showed consistency with the patient responses in all except one patient, it is important to note that the instrument was validated only for the patients. When administered to the caregivers, each element of the survey was modified from its original form, and it is possible that with these modifications the statements do not carry the same meaning as when administered to the adolescent patient. In the event of discrepancy between patient and caregiver responses, each modified question would have been studied through focus groups, and validation would have been performed separately for the caregiver proxy version. Our results are similar to other findings in the literature examining quality of life according to the child compared with the caregiver proxy, where children and caregivers tend to agree on quality-of-life measures looking at physical function: in general, caregiver DIs indicated a slightly lower quality of life for the child, both in our study and in the literature, but without statistical significance.

Our study has several weaknesses. One weakness of the study is that the instrument has been validated only in English. Therefore, for patients who did not have adequate command of the English language, discrepancy may exist in their understanding of some or all of the questionnaire. In addition, the sample size of 56 investigated patients was small. Approximately 30% of those diagnosed as having PVFM did not complete the DI before or after therapy. This was owing to lack of attendance at voice therapy or lack of follow-up visits after therapy. Although this response rate is slightly lower than expected (and it is possible that this may have introduced some selection bias into the study), we believe that the 56 individuals who participated adequately represented the study population. Furthermore, the variability in the number of sessions needed to achieve level 3 among patients may have affected the results. Patients who had more sessions may have achieved lower DIs than those who had fewer sessions, thereby being unable to achieve level 3 and decreased DIs. Last, because this is the first study to date using a uniform measure to assess the severity and treatment response in pediatric PVFM, no criterion standard exists for comparison of outcomes. Therefore, discriminant validity and convergent validity were determined only by comparing responses to a query of global improvement or worsening with treatment.

Conclusions

Future directions include validation of the DI in the caregiver proxy form to account for the slight difference between patient and caregiver responses. In addition, the next step in further understanding PVFM is the identification of pretreatment characteristics to predict the outcome of voice therapy. Previous investigations have identified factors that may affect response to therapy such as asthma, reflux, and underlying anxiety disorders. With the DI having been validated in this pediatric population, it can be used to assess outcomes in patients with additional issues that may contribute to PVFM.


