Efficiency and Cutoff Values of Self-Assessment Instruments on the Impact of a Voice Problem

*,†Mara Behlau, *Glaucya Madazio, *,†Felipe Moreti, *Gisele Oliveira, *Luciana de Moraes Alves dos Santos, *Bruna Rabelo Paulinelli, and *Euro de Barros Couto Junior, *†São Paulo, São Paulo, Brazil

Summary: Objectives. To evaluate the efficiency of four self-assessment questionnaires that rate the impact of a voice problem on the individual's life: Voice-Related Quality of Life (V-RQOL), the original and reduced versions of the Voice Handicap Index (VHI) and VHI-10, Vocal Performance Questionnaire (VPQ), and Voice Symptom Scale (VoiSS).

Methods. Data from 975 subjects, 486 with a diagnosis of dysphonia and 489 vocally healthy individuals, were submitted to the receiver operating characteristic (ROC) curve analysis to obtain the cutoff values that determine the discriminating power of these instruments (presence of dysphonia vs healthy voice).

Results. The ROC curve analysis showed that the most efficient questionnaires were the VoiSS and the VHI. Results showed that they presented as a perfect classification based on their efficiency, specificity, and sensitivity values (all three of them = 1). The VHI-10 and the V-RQOL showed excellent classification (VHI-10: efficiency = 0.991; specificity = 1; sensitivity = 0.981; V-RQOL: efficiency = 0.914; specificity = 0.860; sensitivity = 0.967). Finally, the VPQ showed a good level of classification (efficiency = 0.828; specificity = 0.824; sensitivity = 0.831). The cutoff values for the instruments are as follows: VoiSS = 16 points, VHI = 19 points, VHI-10 = 7.5 points, V-RQOL = 91.25, and VPQ = 20.5 points. These values are important for screening large populations as well as for helping in the decision-making process of clinical management. The cutoff values for maximum sensitivity and specificity of the instruments that did not produce perfect classification are as follows: VHI-10: sensitivity = 5; specificity = 7.5, V-RQOL: sensitivity = 86.25; specificity = 98.75, and VPQ: sensitivity = 15.5; specificity = 31.5.

Conclusions. Both the VoiSS and the VHI are perfect classifiers. The VHI-10 and the V-RQOL are excellent classifiers, and the VPQ is good at discriminating individuals with dysphonia from the ones without dysphonia.

Key Words: Voice–Dysphonia–Self-assessment–Validation studies–Protocols–Speech, Language, and Hearing Sciences.

INTRODUCTION

Voice disorders, also called dysphonias, occur in 3–9% of the population.[1] They affect quality of life in several different ways. Traditionally, the evaluation of patients with voice disorders is a multidimensional process, including at least a laryngeal examination, perceptual, and acoustic analysis.[2] However, the patient's experience of living with dysphonia cannot be inferred directly by these standard clinical assessments. Measuring what patients perceive about their health condition is essential, especially because there is typically a low correlation between the patient's and the clinician's subjective voice analyses.[3,4] Therefore, only the patient can provide real information about his/her experience with the voice problem, which cannot be obtained with any objective analysis.[5,6]

The concept of health evolved remarkably during the last decades after the formulation of the International Classification of Functioning.[7] The concept advanced toward measuring a disability as a difficulty found at least in one of the three functioning domains: impairment, activity limitations, and participation restriction. Disability is a consequence of the interaction between health and contextual factors, whether the latter are environmental or personal. Hence, it is essential for the diagnostic process to understand the perspective of the individual who experiences the problem. The perspective of the patient is usually obtained by means of self-assessment questionnaires that rate the impact of a certain deviation, disorder, or illness.

Numerous self-assessment questionnaires have been developed since the 1990s.[8–12] They were developed not only to quantify the impact of a voice problem and to evaluate the patient's progress but also to contribute in therapeutic management. These instruments became very popular both in the clinical and scientific settings. They achieved fast international popularity never seen before with any other approach, including acoustic analysis.[13] The Voice Handicap Index (VHI)[8] was the first questionnaire introduced in the area. It has a specific purpose of assessing the impact of dysphonia on patients' quality of life. This totally new perspective was proposed during a time when the trend was to use quantitative measures from heavy instrumentation for vocal analysis. Although the several other self-assessment instruments that followed the VHI were built in a somewhat structured manner, their development was based on varied criteria. This is specifically true in regard to the conceptual and empirical basis used for generating the instruments' content, that is, compiling of items. Some instruments included only data from patients' records, others included health professionals' and patients' points of view and/or the scientific literature in the area.

Accepted for publication May 11, 2015.

From the *Centro de Estudos da Voz–CEV, São Paulo, São Paulo, Brazil; and the †Department of Speech-Language Pathology and Audiology, Universidade Federal de São Paulo–UNIFESP, São Paulo, São Paulo, Brazil.

Address correspondence and reprint requests to Mara Behlau, Rua Machado Bittencourt, 361. São Paulo, SP 04044-001, Brazil. E-mail: mbehlau@uol.com.br

Journal of Voice, Vol. 30, No. 4, pp. 506.e9-506.e18
0892-1997/$36.00

http://dx.doi.org/10.1016/j.jvoice.2015.05.022

The numerous dysphonia self-assessment questionnaires available represent conceptual differences, such as focusing on the handicap,[8] the quality of life,[9] the loss of vocal endurance,[10] or even on the identification of vocal symptoms.[12]

These instruments must be submitted to a series of controlled procedures to be used in a language other than the original. Procedures include linguistic and cultural adaptation, validity, reproducibility, responsiveness, and reliability measures.[14]

Many vocal self-assessment questionnaires that were originally developed in English have been already validated into Brazilian Portuguese by the group of authors of this study. Among these instruments are Qualidade de Vida em Voz (QVV)[15] (Voice-Related Quality of Life [V-RQOL]),[9] Índice de Desvantagem Vocal (IDV)[16] (Voice Handicap Index [VHI; Jacobson et al 1999]),[8] and its reduced version (IDV-10[17] e VHI-10),[18] Questionário de Performance Vocal (QPV)[19] (Vocal Performance Questionnaire [VPQ]),[10] and Escala de Sintomas Vocais (ESV)[20] (Voice Symptom Scale [VoiSS]).[12]

As a result of their validation studies, the mean scores of all the aforementioned questionnaires are known for both dysphonic and vocally healthy individuals.[15,16,19,17,20] These values are presented in this article in Table 1. However, as of now, the literature has not presented sufficient information about the cutoff values of these questionnaires that discriminate healthy from disordered individuals and about the degree of classification based on their efficiency. One of the most adequate analysis used for determining the discriminatory power of a binary classification system, that is, its efficiency in the task it was developed to do, is the receiver operating characteristic (ROC) curve, known as ROC curve. This analysis is a statistical procedure originated from the signal detection theory used for measuring the accuracy of sensorial judgments.[21]

The ROC curve represents the relationship between the sensitivity (ability of a test to correctly identify individuals with the problem in question, ie, the ratio of true positives) and the specificity (ability of a test to correctly identify individuals without the problem in question, ie, the ratio of true negatives) of any given test. The ROC curve is a simple analytical procedure for determining the real value from which two categories are discriminated.[22] This analysis considers the highest possible values of sensitivity and specificity, concomitantly combined with the highest values of efficiency (ability of a test to correctly identify both the positive and negative cases, ie, both the presence and absence of the illness or disorder) and product (a value that confirms the efficiency of the test and must accompany the efficiency's values to indicate greater accuracy of results) producing what is called the cutoff value.

The maximum value of 1.0 for sensitivity, specificity, efficiency, and product indicates that the instrument is able to perform a perfect classification as to what it is proposed to evaluate.

The cutoff value is a number from which the result of a test is classified either as positive (presence of deviation, disorder or illness that is being tested) or negative (absence of what is being tested). If the result found is smaller than the cutoff value, the result of a test is classified as negative and vice versa. The ROC curve allows for the comparison of several diagnostic tests, which is one of its most important applications. To determine whether two ROC curves are equal or different, the area under the curve (AUC) is calculated. This area classifies the level of accuracy of a diagnostic test. Consequently, the AUC measures the performance of the test, for instance in the voice area, its accuracy to identify individuals with voice problems. A test that is not able to discriminate between individuals with or without a certain disorder has an AUC of 0.5 (casual identification). Only areas that have values >0.7 are considered satisfactory.

There are only few studies that determined the ROC curve of self-assessment instruments that evaluate the impact of a dysphonia on the individual's life. All these studies have used specifically the original version of VHI. The studies that used the VHI with 30 items showed cutoff values that vary from 12 to 20. All studies found AUCs that were at least satisfactory. The first study[48] compared patients with glottic cancer and benign laryngeal lesions. They used the Dutch version of the questionnaire and found a cutoff value of 15 points, with sensitivity of 0.97 and specificity of 0.86. Two other research groups analyzed the German[23] and the Polish[24] versions of the VHI in patients with several different types of dysphonia. The cutoff value obtained was 12 points for the Polish study with sensitivity of 0.98 and specificity of 0.95. The Swedish version of the VHI[25] obtained a higher cutoff value of 20 points with a sensitivity of 0.77 and a specificity of 0.87. However, the authors highlighted that one of the limitations of the study

TABLE 1.
Characteristics of the Validation Studies With Subjects Distribution and Mean Total Scores of the Questionnaires, According to Chronologic Completion of the Validation

Questionnaire	Authors and Year of Validation Into Brazilian Portuguese	Dysphonic Individuals					Vocally Healthy Individuals				
		Male	Female	Total	Mean Age	Average Score	Male	Female	Total	Mean Age	Average Score
V-RQOL	Gasparini and Behlau 2009[15]	19	95	114	41.3	65.9	31	89	120	43.0	98.0
VHI	Behlau et al 2011[16]	14	38	52	42.3	48.1	20	44	64	41.1	4.5
VPQ	Paulinelli et al 2012[19]	19	141	160	41.0	27	31	134	165	33.0	19
VHI-10	Costa et al 2013[17]	6	54	60	46.9	18.6	6	44	50	43.4	1.7
VoiSS	Moreti et al 2014[20]	56	104	160	43.0	49.4	49	91	140	42.2	7.1