The Applicability of the Dysphonia Severity Index and the Voice Handicap Index in Evaluating Effects of Voice Therapy and Phonosurgery

Marieke M. Hakkesteegt, Michael P. Brocaar, and Marjan H. Wieringa, *The Netherlands*

Summary: The objective was to investigate the applicability of the Dysphonia Severity Index (DSI) and the Voice Handicap Index (VHI) in evaluating effects of intervention between groups of patients and for intrasubject differences and whether DSI and VHI are complementing measurements. Analyses of measurement data before and after intervention of 171 patients with voice disorders. The voice quality was measured objectively with the DSI. The perceived voice handicap was measured with the VHI. Three groups of patients were used: patients who had voice therapy, phonosurgery, or no intervention. DSI and VHI improved significantly after intervention in the voice therapy and the surgery group (median difference DSI 1.19 and 3.03, VHI −8 and −26, respectively). The intrasubject results were analyzed based on the test-retest variability of DSI and VHI. Significant better DSI and VHI scores after intervention were found in, respectively, 22% and 38% of the patients with voice therapy, and 56% and 78% of the patients with surgery. In the no intervention group, this was 11% and 12%. In 37% of the patients, the differences before and after intervention in DSI and VHI were in discordance. The DSI and VHI are able to show significant differences after intervention for voice disorders between groups of patients. The DSI and VHI can be used to determine a significant intrasubject result of intervention. The DSI and VHI measure each different aspects of the voice and are complementing measurements. The DSI is therefore applicable in clinical practice for objective evaluation of voice quality and the VHI for subjective evaluation of the perceived handicap by the patient self.

Key Words: Voice disorders–Voice quality–Treatment outcome–Dysphonia Severity Index–Voice Handicap Index.

INTRODUCTION

Evaluating effects of intervention is of growing importance in today's health care, because of the need for evidence-based intervention. For voice disorders also, research on the effects of intervention is needed. There are however not yet well-accepted standardized instruments that can be used to assess the effects of intervention for voice disorders. When evaluating effects of intervention, there are two different aspects to take into account: the differences between groups of patients (intersubject differences) and the difference within one patient before and after intervention (intrasubject differences). The differences in outcome between groups are needed for research purposes: to compare a new type of intervention with a commonly used type of intervention, or to determine what the best type of intervention is for a certain diagnosis. Therefore, the intersubject variance of the used measurements has to be known. In daily clinical practice, it is important to be able to interpret differences between measurements of one patient made on different points in time (eg, before and after intervention). To know whether differences are significant, the intrasubject variance of the used measurements has to be known.

Because voice disorders consist of different aspects (voice quality, voice handicap), several measurements should be used. Clinical assessment of voice disorders should consist of (video)laryngostroboscopy, perceptual voice assessment, objective measurements (acoustic analysis and aerodynamic measurements), and subjective self-evaluation of voice.[1] However, not all these aspects appear equally suitable for evaluating effects of intervention.

Although (video)laryngostroboscopy is a very important clinical tool for diagnosing and evaluating patients with voice disorders, it has not been widely used as a research tool because the interpretation is subjective and reliable and quantifiable tools for research purposes are not yet available.[2] For the perceptual voice assessment, the "GRBAS scale" as introduced by Hirano[3] is widely used. The reliability of grade has been investigated.[4,5] However, these investigations are expressed as levels of agreement (kappa values) for inter- and intra-rater and test-retest reliabilities, and are not expressed as the intrasubject variance. The reason for this is probably that grade is scored on a categorical scale and consequently calculations cannot be made. Therefore, grade appears not to be suitable for evaluation of intervention effects, neither between groups of patients, nor for intrasubject differences.[6,7] There is no consensus on what objective measurements for voice quality are best suitable to measure effects of intervention. In studies describing effects of intervention for voice disorders, a variety of measurements are used. The choice for the used measurements in evaluation studies can be based on expected changes in specific aspects of voice quality.[8–11] However, in daily clinical practice it is most practical to use the same objective measurement for all voice disorders. It is already known that multiparametric measures are more suitable for evaluation of voice quality than single measures.[12–15] The Dysphonia Severity Index (DSI)[14] is such a multiparametric measure. The DSI has a good relationship with the perceptual evaluation on grade of the GRBAS scale.[16] An advantage of the DSI is that the

Accepted for publication June 30, 2008.

From the Department of Otorhinolaryngology and Head and Neck surgery, Erasmus MC – University Medical Center Rotterdam, The Netherlands.

Address correspondence and reprint requests to Marieke M. Hakkesteegt, Department of Otorhinolaryngology and Head and Neck surgery, Erasmus MC – University Medical Center Rotterdam, P.O. box 2040, 3000 CA Rotterdam, The Netherlands. E-mail: m.hakkesteegt@erasmusmc.nl

Journal of Voice, Vol. 24, No. 2, pp. 199-205
0892-1997/$36.00

doi:10.1016/j.jvoice.2008.06.007

parameters can be obtained relatively quick and easy by speech pathologists in daily clinical practice.

For the self-evaluation of voice, the Voice Handicap Index (VHI) is a widely used tool.[17] The VHI is a subjective self-administered questionnaire addressing the patients perceived disability. The VHI and the DSI measure each different aspects of voice and the outcomes on both measurements are therefore not necessarily related. The patient's perception of the voice disorder is not only related to voice quality (as measured with the DSI) but is also related to, for example, professional and social vocal demands and personal aspects.

For both the DSI and VHI, the clinical significance (standard deviation [SD]) is known, obtained from test-retest variability investigations.[18,19] This clinical significance is necessary to interpret the differences between measurements before and after intervention. Therefore, these measurements could be suitable to evaluate effects of intervention. The VHI is already used for evaluation, the DSI however is until now only used on a very limited scale.

We hypothesized that it is possible to evaluate results of intervention for voice disorders with a widely applicable objective measurement. We also hypothesized that measurements of different aspects of voice disorders will complement each other. The purposes of this study were as follows:

To investigate whether the DSI and the VHI can be used to evaluate effects of different types of intervention for voice disorders between groups of patients.
To investigate whether the DSI and VHI can be used in daily clinical practice to determine a significant intrasubject effect of intervention for voice disorders.
To investigate whether the DSI and VHI are complementing measurements, by investigating the relationship between the differences before and after intervention on the DSI and the VHI.

MATERIALS AND METHODS

Subjects

The measurements of the patients with voice disorders visiting the outpatient clinic of our department of Otorhinolaryngology are collected in a database. The measurements are done at the first visit and at follow-up visits. When patients have voice therapy, the follow-up measurements are done at least 3 months after the first therapy session. When patients have phonosurgery, the follow-up measurements are done at least 2 months after surgery. When there is no intervention, the minimum time interval had to be 6 weeks. From this database, the data of the patients of whom DSI measurements were available of their first visit (preintervention) and at least one follow-up visit (postintervention) were used. In case of several follow-up measurements, the last measurement was used as postintervention measurement. This resulted in measurement data of 171 patients (74 male, 97 female) with a mean age of 43 years (range, 15–82, SD 15 years). From 122 of those patients also, VHI measurements of both visits were available. The VHI was not for all patients available, partly because we introduced the VHI later than we started with DSI measurements, partly because some of the patients were not able to complete the questionnaire due to language problems. The median time interval between the measurement of the first and the last visit was 29 weeks (minimum 6, maximum 171 weeks). Patients were classified in three diagnosis groups: nonorganic dysphonia, mass lesions (nodules, polyps, cysts, laryngitis, and edema), and paresis/paralysis (unilateral and bilateral paresis and paralysis). This classification was already used in a previous study.[16] There were two intervention groups: surgery combined with voice therapy (further called "surgery group") and voice therapy only. Patients who visited the department only for follow-up, while there was no (further) intervention were assigned to a "no intervention" group. These were patients for whom no intervention was available or who chose not to be treated.

Procedures

The patients were asked to fill in a VHI form while they were in the waiting room. All patients were examined by one of the two speech pathologists of the department, who measured the DSI parameters. Thereafter, the clinical diagnosis was made with laryngostroboscopy by one of the two ear, nose, and throat (ENT)/voice specialists of the department.

Measurements

The parameters used for DSI measurements are the highest fundamental frequency (F_0-high in Hz), lowest intensity (I-low in dB sound pressure level (SPL)), maximum phonation time (MPT in s), and jitter (%). The DSI is constructed as $\text{DSI} = 0.13 \times \text{MPT} + 0.0053 \times F_0\text{-high} - 0.26 \times I\text{-low} - 1.18 \times \text{Jitter (\%)} + 12.4$. It is constructed such that a perceptually normal voice (grade 0) corresponds with a DSI of +5; a severely dysphonic voice (grade 3) corresponds with a DSI of −5. Scores beyond this range (higher than +5 or lower than −5) are also possible. To obtain I-low, the subjects were asked to phonate an /a/ as softly as possible at a comfortable pitch. To obtain F_0-high, they were asked to produce an /a/, starting at a comfortable pitch going up to the highest and down to the lowest pitch. This instruction was accompanied by a demonstration by the speech pathologist. To measure MPT, the subjects were asked to inhale deeply and sustain an /a/ for as long as possible at a comfortable pitch and loudness. The MPT was recorded three times; the longest measured phonation time in seconds was used. To calculate jitter, the subjects phonated three times an /a/ at a comfortable pitch and loudness during approximately 3 seconds. The jitter was calculated on a sample of 1 second, starting half a second after the voice onset. The lowest result of the three calculations was used.

Equipment

Intensity and frequency measurements were obtained with an automatically recording phonetograph (Pabon/Laryngograph 1997). The *Multi-Speech* program (Kay Elemetrics, Lincoln Park, NJ) was used for calculating jitter. Audio recordings were made with a sampling rate of 11,025 Hz and 16 bits quantization. A Sennheiser microphone (BG 2.0 dyn) was used. The