Cross-cultural adaption and validation of the Danish Voice Handicap Index-10

Running title: Danish VHI-10

MD Line Schiøtt Nissen\textsuperscript{1α}, MD Joyce Schultz\textsuperscript{1α}, MD Jacob Galili\textsuperscript{1}, MA Trine Printz\textsuperscript{2}, MD Camilla Slot Mehlum\textsuperscript{2}, MD Ågot Møller Grøntved\textsuperscript{2}, PhD Jesper Roed Sorensen\textsuperscript{1,2}

1) Department of Otorhinolaryngology, Lillebaelt Hospital, Kabeltoft 25, DK-7100 Vejle, Denmark
2) Department of ORL Head & Neck Surgery and Audiology Odense University Hospital, J.B. Winsløws vej 4, DK-5000 Odense C, Denmark

α Shared first authorship

Abstract: 288 words
Manuscript: 2405 words
Figures: 1
Tables: 2

Keywords: Voice Handicap Index-10, vocal, validity, reliability, dysphonia, Danish

Conflicts of interest statement:

The authors have no conflicts of interest. The study has been funded by the Region of Southern Denmark, the main employer of the authors.

Corresponding author
Jesper Roed Sorensen, MD, ORCID: 0000-000-4351-8873
Department of Otorhinolaryngology
Lillebaelt Hospital
Kabeltoft 25, DK-7100 Vejle, Denmark
E-mail: jesper.roed.sorensen@rsyd.dk. Phone: +45 65 41 28 10. Fax: +45 66 19 26 15

Abstract
Background: The Voice Handicap Index 30 (VHI-30) is a much-used voice specific quality of life questionnaire. A shortened ten item version has been developed by eliminating redundant items using item analyses. This is the first Danish translation of the Voice Handicap Index 10 (VHI-10).

Objectives: To evaluate the psychometric properties of the Danish VHI-10 questionnaire.

Study design: Cross-sectional survey study.

Methods: A Danish translation of the VHI-10 was answered by 72 patients with voice disorders of different etiology (neurogenic, functional, and structural) and by a control group of 94 vocally healthy individuals. Thirty-two patients and 68 controls participated in a test-retest reliability analysis. The internal consistency, test-retest reliability, and clinical validity were assessed.

Results: Excellent internal consistency was found in the patient group with a Cronbach’s α of >0.90. In the control group the internal consistency was good with a Cronbach’s α of 0.88. Test-retest reliability was good with intra class correlation coefficient of 0.94 (95% confidence interval [95%CI]: 0.88–0.97) for patients and 0.82 (95%CI: 0.73-0.89) for the control group. This indicates a sufficient reliability of the questionnaire. The correlation between the Danish VHI-10 score and the patient’s perception of the severity of the voice disorder was 0.75 (p<0.001) indicating good clinical validity of the Danish VHI-10.

Conclusion: The newly translated Danish VHI-10 was validated and performs similar to the original VHI-10. It showed good internal consistency, test-retest reliability, and clinical validity. The questionnaire is preferably for use in patients with moderate to severe voice complaints as its ability to distinguish mild voice changes from healthy voices is limited. However, the questionnaire is capable of assessing patients’ perception of the severity of their voice disorder and is available for use in daily practice and in research projects.
1.0 Introduction

Voice disorders can have a significant impact on the quality of life as shown by Health-Related Quality of Life (HRQoL) questionnaires[1, 2]. The impact on HRQoL of voice disorders is comparable to, or even larger than that of other diseases such as angina, sciatica, and chronic rhinosinusitis[1, 2]. The self-perceived severity of a voice disorder provides important information to clinicians in daily clinical practice[3]. It can be assessed by questionnaires such as the Voice-Related Quality of Life (V-RQOL)[4] or the Voice Handicap Index (VHI) questionnaires[5].

In 1997 Jacobsen et al. developed a 30-statement questionnaire, the VHI-30[5], which has been validated and translated into multiple languages[6-11]. Subsequently, item analyses has identified the ten most robust items of the VHI-30 and these have been selected for a shorter VHI-10 questionnaire[12], which is less time-consuming for patients. The VHI-10 has shown a good correlation to the results of the VHI-30 and has been thoroughly validated by internal consistency, reliability, and test-retest reliability tests[12, 13]. In addition, normative values[14] and minimal important differences[15] has been established for the questionnaire and it now exists in multiple languages[16-18]. Thus, the VHI-10 is a well-suited clinical tool for patient self-evaluation of the severity of a voice disorder. However, a Danish translation has not yet been constructed and validated. The aim of this study was to evaluate the psychometric properties of a Danish version of the VHI-10 questionnaire.

2.0 Materials and Methods

2.1 Scale content and translation

The VHI-10 consists of ten items covering different subjective aspects of a voice disorder. The items are rated on a five-point Likert scale from 0 (never) to 4 (always). The score ranges from 0 to
40 points with a higher score denoting greater severity of the patient’s own perception of his or her voice problem. The Danish version of the VHI-10 was constructed according to the standards of the World Health Organization for cross-cultural translations of questionnaires [19]. The cultural term voice handicap or voice impairment exists as a phenomenon in Denmark implying voice related problems relevant for a person’s daily life. Initially, forward translation was performed of the original American VHI-10 into Danish by a bilingual translator and a health professional with Danish as the native language. The translated Danish VHI-10 was evaluated by a bilingual expert panel of laryngologists and speech-language-pathologists and a revised version was back-translated by an independent bilingual translator, whose native language was English and who had no prior knowledge of the questionnaire. The original American VHI-10 was compared to the back-translated version of the Danish VHI-10 and the process was reiterated three times until an acceptable version was reached. The proposed final version was pilot tested on a target population and slightly amended after interviewing participants and lastly a report of the process was written [20]. In the process, linguistic challenges was met with patient’s understanding of the item “My voice problem causes me to loose income”, which led to the multiple reiterations. This issue may be related to the significant differences between the organization of health care, disability care, and system for sickness benefits between the US and Denmark, as in Denmark temporary voice problems rarely leads people to loose income.

2.2 Participants

Patients with voice disorders were included prospectively and they answered the Danish VHI-10 questionnaire at their initial contact with one of our two laryngological out-patient clinics. Based on etiology, the patient group was separated into three subgroups: Group 1) Neurogenic Voice
Disorder including Recurrent laryngeal nerve paresis; Group 2) Functional Voice Disorders including hypokinetic dysphonia and muscle tension dysphonia; and Group 3) Structural Voice Disorders including intracordal cyst, polyp, nodules, phono trauma, papillomas, vocal fold atrophy, Reinke’s edema, leukoplakia, and early laryngeal cancer. Only patients that were expected to have a stable level of voice complaints for the following 2-4 weeks were considered for inclusion. A control group of vocally healthy individuals was recruited from hospital staff, students, and relatives of both patients and staff. Suggested controls were not included if they had self-reported voice complaints at present or a previous voice disorder requiring intervention.

Self-reported subjective severity of voice complaints was answered on a 7-point Likert scale ranging from no complaints (1 point) to severe complaints (7 points). Scores were grouped into three levels of voice complaints: 1-3 points (none or mild complaints), 4-5 points (moderate complaints), and 6-7 points (severe complaints).

Test-retest reliability of the Danish VHI-10 questionnaire was assessed using repeated responses from patients and controls that completed the questionnaire twice with approximately two weeks interval. Patients were excluded if they at the time of re-test reported a significant change in level of voice complaints since the first test (supplementary question to the Danish VHI-10 questionnaire: “My voice is significantly changed since my last reply to the Voice Handicap Index-10 questionnaire: yes/no”).

This study was approved by the regional ethics committee. All participants signed an informed consent declaration before participation.

2.3 Statistical analyses
The internal consistency was estimated by the Cronbach’s α coefficient. The consistency was considered sufficient for α >0.7, good for α >0.8, and excellent for α >0.9[21].

Clinical validity was assessed by the Spearman’s rank correlation coefficient (SRCC) investigating the association between VHI-10 score and self-reported voice complaints. SRCC values of 0.60-0.80 were considered a strong correlation and values >0.80 as a very strong correlation.

For test-retest reliability analyses the intraclass correlation coefficient (ICC) was used to assess the proportion of the variability from between-subject differences out of the total variability. ICC values ≤ 0.50 were considered poor reliability, values of 0.51-0.74 as moderate reliability, and values ≥ 0.75 as good reliability[22].

The ability of the VHI-10 to distinguish between the three etiological subgroups of patients with voice disorders and the controls, was assessed by the Kruskal-Wallis and the Conover-Iman pairwise comparison test. The impact of age, sex, and having a voice disorder on the individual and total VHI-10 score was assessed by multiple linear regression analysis.

### 3.0 Results

The 72 patients that completed the Danish VHI-10 (mean age 59 ± 15 years, 53% females) were divided into the three subgroups Neurogenic Voice Disorders (n=12 [17%]), Functional Voice Disorders (n=39, [54%]), and Structural Voice Disorders (n=21 [29%]). Forty-two patients completed the follow-up questionnaire of which ten patients, all with Functional Voice Disorders, were excluded due to self-reported significantly changed voice complaints (independent supplementary question) since their first response. These patients also had notable differences in Danish VHI-10 scores from the first to the second response (mean improvement 13.2±8.0 points).

The control group consisted of 94 individuals (mean age 40 ± 12 years, 74% females) of which 68
completed the follow-up questionnaire. The patients and controls differed significantly in age (p<0.001).

The self-reported severity of voice complaints are summarized in Table 1. None of the controls reported moderate or severe voice complaints. Multiple regression analyses showed that the Danish VHI-10 score was not influenced by age (p=0.99) nor sex (p=0.19), but the presence of a voice disorder significantly impacted the VHI-10 score (p<0.001).

The mean VHI-10 score for the Danish patients was 18.1 ± 9.4 (range 2-39) compared to 1.2 ± 2.3 (range 0-13) for the control group (p<0.001) (Table 1). The internal consistency for the VHI-10 questionnaire was good to excellent with overall Cronbach’s α value of 0.92 for the patient group and 0.84 for the control group (Table 2).

The clinical validity of the Danish VHI-10 was also good with SRCC value of 0.75 (p<0.001), indicating a strong correlation between VHI-10-score and self-reported voice complaints. ICC scores (test-retest reliability) were good with 0.94 (95%CI: 0.88–0.97) in the patient group and 0.82 (95%CI: 0.73-0.89) in the control group with only minor differences of 1.0 ± 3.0 for patients (p=0.11) and -0.4 ± 1.5 for controls (p=0.06) between the first and second test.

When using the Kruskal-Wallis and the post-hoc Conover-Imam testing, the VHI-10 score of each of the three etiological subgroups of patients was significantly different from the control group (p<0.001). None of the etiological subgroups differed significantly from each other (Figure 1).

### 4.0 Discussion

The Danish VHI-10 showed good internal consistency, test-retest reliability, and clinical validity. This was comparable to the original VHI-10 questionnaire and in alignment with existing normative
The internal consistency of the questionnaire was good to excellent; indicating that this version covers various subjective aspects of voice complaints. Previously, normative values for voice healthy individuals of the original American questionnaire have been established with a VHI-10 score of mean 2.8 ± 3.9 based on 158 participants[14] comparable to our 94 voice healthy individuals (mean 1.2 ± 2.3). The study established a cut-off level for possible voice disorders at a mean VHI-10 score plus two standard deviations above the normal voice implying a value above 10.7 points as indicative of a voice disorder [14]. The mean score in our patient group was well above this level (18.1 ± 9.4 points). However, 20 of our patients (27%), who were all referred to our laryngological clinics because of a voice disorder/impairment, scored 11 points or less. This may have several explanations: (1) it might suggest cross-cultural differences in the understanding of mild voice impairment, (2) it might reflect tax-based and easy access health care services, which potential could lead to over-use, (3) it could reflect difficulties in the VHI-10 questionnaire in identifying patients with mild voice complaints and distinguishing them from voice healthy individuals. The Danish VHI-10 questionnaire is thus probably insufficient for tracking patients with subtle voice changes, as it is unable to distinguish between patients with minor voice disorders and participants with no voice complaints.

The test-retest reliability was excellent for patients and good for controls indicating a reliable questionnaire, suitable for longitudinal studies. The patient group and the control group had changes in scores of 1.0 ± 3.0 and -0.4 ± 1.5 points respectively in the test-retest interval which is well below the estimated minimal important measurement difference of six points of the original VHI-10 score [15]. Misono et al. described limitations of the original VHI-10’s ability to distinguish between no change and a small change in the perception of the severity of the voice disorders, and indicated that the questionnaire might not be suitable for evaluating small differences after treatment[15]. The results of our study indicate that a change in the Danish VHI-10 score after any
intervention most likely is caused by the intervention and not by random variation. All ten patients excluded from the test-retest analyses had subjective voice changes and VHI-10 scores of more than this estimated minimal important measurement difference level of six points [15].

The strong correlation between VHI-10 score and self-reported severity of voice disorder indicated a robust clinical validity. However, the correlation of 0.75 implies that the Danish VHI-10 questionnaire cannot be used to distinguish patients with voice disorders from healthy persons, nor is that the purpose of the original questionnaire [5, 12]. The questionnaire cannot be used as a voice screener, as the reliability of the questionnaire in distinguishing between mild voice complaints and voice healthy individuals are inadequate. However, it is suitable for estimating the subjective severity of patients with moderate to severe voice complaints. The results of the validation show that the Danish VHI-10 has an acceptable quality for use in patients with mild to severe voice disorders in Denmark. It is a well-functioning questionnaire for assessment of the patients’ own perception of their voice disorders. The quality of the newly validated Danish VHI-10 is comparable to the existing validated Danish VHI-30, however, with a slightly lower, but still excellent Cronbach’s α of 0.92 vs. 0.97. This is a consequence of fewer questions in the VHI-10. However, the Danish VHI-10 performs better than the Danish VHI-30 in both clinical validity (0.75 vs. 0.69) and in test-retest reliability with ICC scores of 0.94 vs. 0.93 for patients and 0.84 vs 0.78 for controls, indicating superior performance of the VHI-10.

Some limitations of our study also need to be addressed. Firstly, the distribution of patients in the etiological subgroups of voice disorders was not evenly distributed and the patients with Neurogenic Voice Disorders were all diagnosed with recurrent laryngeal nerve paralyses. This makes it necessary to be cautious in generalizing the conclusions from this patient group to other patients with other Neurogenic Voice Disorders. This reflects, however, the referral pattern in our institutions. Secondly, the recruitment of the controls led to a significantly younger population in
the control group compared to the patient group, which may confound the analyses. The VHI-10 scores in the patient group ranged from 2 points to 39 points. This large variation may possibly be explained by the various underlying etiologies, ranging from minor functional voice disorders to potentially severe voice disorders as early glottis cancers or recurrent laryngeal nerve paralyses. One single individual in the control group had a VHI-10 score of 13 points, which is above the 11 point cut-off value established in the normative values [14]. This might indicate an underlying undiagnosed voice disorder, but this remains speculation as endoscopy or voice assessment was not performed in the control group. However, this does not seem to be important for the conclusions of the study. Lastly, despite the wording of the questionnaire having been carefully translated, back-translated, and amended after additional test, the translation of an original American questionnaire into a Danish context might not include all aspects of cultural differences in the meaning of a voice disorder. However, we did not find any evidence of a weakness here as the patients’ subjective evaluations of the voice disorders correlated well with the Danish VHI-10 score.

5.0 Conclusions

The VHI-10 translated into Danish has been thoroughly validated and found to be comparable to the original VHI-10. It has a good internal consistency, test-retest reliability, and clinical validity. The questionnaire is preferably for use in patients with moderate to severe voice complaints as its ability to distinguish mild voice impairments from healthy voices is limited. As a consequence, the questionnaire cannot be used for screening purposes to identify patients with minor voice complaints. However, this has never been the intention of the questionnaire. This version is available for future use in daily practice in voice clinics and in research projects.
6.0 Acknowledgments

Simone Lind and Mia Egsgaard, speech language pathologists, are acknowledged for the help and assistance in translating the American VHI-10 questionnaire into Danish. Sille Nielsen and Laura Lily Hauerslev students of speech language pathology are acknowledged for their effort in recruiting a part of the control group.
Fig 1. Box plot of the VHI-10 scores for patients with voice disorders, divided according to etiology, and the control group. (*) The results of the control group differed significantly from the three groups of patients (p<0.001).
<table>
<thead>
<tr>
<th>Patient-reported severity of voice complaints</th>
<th>Control group (n=94)</th>
<th>Patient group (n=72)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None/ Mild</td>
<td>94/94, 100%</td>
<td>26/72 (36%)</td>
</tr>
<tr>
<td>None/ mild</td>
<td>29/72 (40%)</td>
<td>17/72 (24%)</td>
</tr>
<tr>
<td>Moderate</td>
<td>72/72 (100%)</td>
<td></td>
</tr>
<tr>
<td>VHI-10 (mean ± SD)</td>
<td>1.2 ± 2.3</td>
<td>11.0 ± 6.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18.2 ±6.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>29.0 ± 7.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18.1 ± 9.4</td>
</tr>
</tbody>
</table>

Voice Handicap Index (VHI-10) and Standard deviation (SD)
<table>
<thead>
<tr>
<th>VHI-10 item (English)</th>
<th>VHI-10 item (Danish)</th>
<th>Patients</th>
<th>Control group</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>My voice makes it difficult for people to hear me</td>
<td>Min stemme gør det svært for andre at høre mig</td>
<td>0.91</td>
<td>0.82</td>
<td>0.96</td>
</tr>
<tr>
<td>People have difficulty hearing me in a noisy room</td>
<td>Andre har svært ved at forstå mig i et støjfyldt lokale</td>
<td>0.91</td>
<td>0.80</td>
<td>0.96</td>
</tr>
<tr>
<td>My voice difficulties restrict my personal and social life</td>
<td>Mine stemmeproblemer begrænser mig i mit personlige og sociale liv</td>
<td>0.91</td>
<td>0.82</td>
<td>0.96</td>
</tr>
<tr>
<td>I feel left out of conversation because of my voice</td>
<td>Jeg føler mig udelukket fra samtaler pga. min stemme</td>
<td>0.90</td>
<td>0.83</td>
<td>0.95</td>
</tr>
<tr>
<td>My voice problems causes me to loose income</td>
<td>Mine stemmeproblemer gør, at jeg tjener mindre end jeg ellers ville</td>
<td>0.92</td>
<td>--</td>
<td>0.97</td>
</tr>
<tr>
<td>I feel as though I have to strain to produce a voice</td>
<td>Jeg føler, at jeg skal anstrenge mig for at frembringe stemme</td>
<td>0.90</td>
<td>0.81</td>
<td>0.95</td>
</tr>
<tr>
<td>The clarity of my voice is unpredictable</td>
<td>Det er uforudsigeligt, hvornår min stemme lyder klar</td>
<td>0.92</td>
<td>0.81</td>
<td>0.96</td>
</tr>
<tr>
<td>My voice problems upsets me</td>
<td>Mine stemmeproblemer gør mig frustreret og ked af det</td>
<td>0.90</td>
<td>0.84</td>
<td>0.95</td>
</tr>
<tr>
<td>My voice makes me feel handicapped</td>
<td>Jeg føler mig handicappet pga. min stemme</td>
<td>0.90</td>
<td>--</td>
<td>0.95</td>
</tr>
<tr>
<td>People ask, &quot;What’s wrong with your voice?&quot;</td>
<td>Folk spørger: &quot;hvad er der galt med din stemme?&quot;</td>
<td>0.91</td>
<td>--</td>
<td>0.96</td>
</tr>
<tr>
<td>Overall</td>
<td></td>
<td>0.92</td>
<td>0.84</td>
<td>0.96</td>
</tr>
</tbody>
</table>

Voice Handicap Index-10 (VHI-10): All controls replied “Never” for the item in question (--).
6.0 References


September 10, 2019.

20. Egsgaard MH, Lind S. *Will I be able to speak again? - Speech therapeutic rehabilitation after total laryngectomy with insertion of voice prosthesis* [Bachelor project]: Department of language and communication, University of Southern Denmark; 2016.
