e-ISSN: 2987-131X

Sriwijaya Journal of Otorhinolaryngology (SJORL)

Journal website: https://phlox.or.id/index.php/sjorl

Development and Validation of the Indonesian Voice Handicap Index Adapted for Javanese and Sundanese Speakers (I-VHI-JS)

Hasrita Soleiman[1*], Andhika Kurnianta Kusuma[2], Eduardo Michael Perez[3]

[1]Department of Otorhinolaryngology, Phlox Institute, Palembang, Indonesia

[2]Department of Oral Health and Dentistry, Phlox Institute, Palembang, Indonesia

[3]Department of Oral Biology, Cobija State Hospital, Cobija, Bolivia

ARTICLE INFO

Keywords:
Indonesian
Javanese
Sundanese
Voice disorders
Voice handicap index

***Corresponding author:**
Hasrita Soleiman

E-mail address:
hasrita.soleiman@phlox.or.id

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.59345/sjorl.v2i2.186

A B S T R A C T

Introduction: Voice disorders significantly impact quality of life. The Voice Handicap Index (VHI) is a widely used patient-reported outcome measure, but its applicability is limited by linguistic and cultural factors. Indonesia, with its diverse languages, requires culturally adapted versions. This study aimed to develop and validate the Indonesian Voice Handicap Index adapted for Javanese and Sundanese speakers (I-VHI-JS). **Methods:** The original English VHI was translated into Indonesian, Javanese, and Sundanese using a rigorous forward-backward translation process. A cross-sectional study was conducted involving three groups: (1) individuals with voice disorders (n=150; Javanese speakers = 75, Sundanese speakers = 75), (2) age- and gender-matched vocally healthy controls (n=150; Javanese speakers = 75, Sundanese speakers = 75), and (3) a test-retest reliability group (n=50; Javanese speakers = 25, Sundanese speakers = 25) from the voice disorder group. Participants completed the appropriate I-VHI-JS version. Otolaryngological examination and acoustic voice analysis (jitter, shimmer, noise-to-harmonics ratio) were performed. Internal consistency (Cronbach's alpha), test-retest reliability (intraclass correlation coefficient - ICC), construct validity (known-groups comparison), concurrent validity (correlation with acoustic parameters), and discriminant validity (receiver operating characteristic - ROC curve analysis) were assessed. **Results:** The I-VHI-JS demonstrated excellent internal consistency (Cronbach's alpha > 0.90 for all versions) and test-retest reliability (ICC > 0.85 for all versions). Significant differences in I-VHI-JS scores were found between the voice disorder and control groups ($p < 0.001$) for all language versions, confirming construct validity. Moderate correlations were observed between I-VHI-JS scores and some acoustic parameters (e.g., jitter, $r = 0.45$, $p < 0.01$; shimmer, $r = 0.40$, $p < 0.01$). ROC curve analysis showed excellent discriminant ability (area under the curve - AUC > 0.80 for all versions). **Conclusion:** The I-VHI-JS is a reliable and valid instrument for assessing voice-related handicaps in Indonesian, Javanese, and Sundanese speakers. It can be used in clinical practice and research to evaluate the impact of voice disorders and monitor treatment outcomes.

1. Introduction

Voice disorders, characterized by deviations in vocal quality, pitch, loudness, or resonance, represent a significant health concern with the potential to profoundly disrupt communication and diminish an individual's overall quality of life. The ramifications of these disorders extend beyond mere vocal impairment, often impacting social interactions, professional efficacy, and psychological well-being. The prevalence of voice disorders is not uniform across populations; however, research indicates that a substantial proportion of individuals will encounter voice-related problems throughout their lives. Certain occupational groups, including teachers, singers, and call center operators, face an elevated risk of developing voice disorders due to the inherent demands placed on their

vocal apparatus. The precise and thorough assessment of voice disorders is paramount for accurate diagnosis, the formulation of effective treatment strategies, and the diligent monitoring of therapeutic progress. While objective measures, such as acoustic analysis and laryngeal examinations, furnish valuable insights into the physiological mechanisms of voice production, they frequently fall short of capturing the subjective experience of the individual grappling with a voice disorder. Consequently, patient-reported outcome measures (PROMs) have emerged as indispensable tools in healthcare, providing a structured framework for evaluating the impact of health conditions from the patient's vantage point.[1-3]

Among the various PROMs available, the Voice Handicap Index (VHI) stands out as a widely accepted and rigorously validated instrument. It is specifically designed to assess the self-perceived consequences of voice disorders across psychosocial, functional, and physical domains. The original VHI comprises 30 items, organized into three distinct subscales: Functional (F), Physical (P), and Emotional (E). In recognition of the need for brevity in certain clinical and research settings, a shortened version, known as the VHI-10, has also been developed and subjected to validation. The VHI has demonstrated remarkable cross-cultural applicability, evidenced by its translation and validation in a multitude of languages. This widespread adaptation underscores the instrument's utility in diverse linguistic and cultural contexts. It is crucial to acknowledge, however, that the process of translation necessitates careful consideration of cultural nuances. Direct, literal translation, without appropriate cultural adaptation, can jeopardize the validity and reliability of the instrument, potentially leading to inaccurate assessments and misinterpretations of patient experiences.[4-6]

Indonesia, the world's fourth most populous country, is characterized by its rich tapestry of linguistic and cultural diversity. While Bahasa Indonesia serves as the national language, the archipelago is home to hundreds of regional languages, reflecting the country's diverse ethnic and cultural groups. Among these, Javanese and Sundanese are two of the most prevalent, each spoken by tens of millions of individuals. The substantial cultural differences that exist between these groups and Western populations, coupled with linguistic variations, underscore the necessity for culturally adapted versions of instruments like the VHI. The use of a generic Indonesian version of the VHI may not be suitable for Javanese and Sundanese speakers due to variations in vocabulary, idiomatic expressions, and cultural conceptualizations of voice and disability. Previous research has consistently emphasized the importance of cultural adaptation in the development and implementation of PROMs. Failure to adequately account for cultural differences can introduce inaccuracies into assessments and lead to misinterpretations of patient experiences. For instance, the expression and understanding of concepts related to voice handicap may vary significantly across cultures. Furthermore, cultural norms and beliefs can exert a powerful influence on how individuals perceive and report their symptoms.[7-10] In light of these considerations, this study was designed to address the existing gap by developing and validating the Indonesian Voice Handicap Index specifically adapted for Javanese and Sundanese speakers (I-VHI-JS).

2. Methods

A cross-sectional study design was the framework for this investigation. Participants were recruited from the otorhinolaryngology outpatient clinics of three major hospitals located in Yogyakarta, Bandung, and Surabaya, cities strategically chosen to represent areas with substantial populations of Javanese and Sundanese speakers. The study population comprised three distinct groups; a Voice Disorder Group; a Control Group; and a Test-Retest Reliability Group. The Voice Disorder Group consisted of individuals aged 18 years or older who had received a diagnosis of a voice disorder by an otorhinolaryngologist. The diagnoses encompassed a range of voice pathologies, including vocal fold nodules, polyps, cysts, muscle tension dysphonia, and vocal fold paralysis. Individuals were excluded from this group if they

presented with cognitive impairment that could hinder their ability to comprehend and complete the questionnaire, had a history of laryngeal surgery within the six months preceding the study, or had other significant medical conditions that might have an impact on their voice. The Control Group was composed of individuals who were age- and gender-matched to the participants in the Voice Disorder Group and had no history of voice disorders. These participants were recruited from the general community. The exclusion criteria for the Control Group mirrored those applied to the Voice Disorder Group. The Test-Retest Reliability Group was a subset of participants drawn from the Voice Disorder Group. These individuals were recruited to complete the I-VHI-JS on two separate occasions, with a two-week interval between the initial assessment and the subsequent reassessment. This procedure was designed to evaluate the test-retest reliability of the instrument. A sample size calculation was conducted to determine the appropriate number of participants needed for the study. This calculation was based on findings from previous VHI validation studies and recommendations for factor analysis. To ascertain test-retest reliability, it was assumed that there would be an expected correlation of 0.80 between the test and retest administrations. With a desired power of 0.80 and an alpha level of 0.05, the minimum required sample size for the test-retest reliability group was determined to be 25 participants. For the assessment of construct validity, assuming a moderate effect size (Cohen's d = 0.5) between the Voice Disorder Group and the Control Group, a power of 0.80, and an alpha level of 0.05, the minimum required sample size for each group was calculated to be 64 participants. To account for potential participant attrition and to ensure adequate representation of both Javanese and Sundanese speakers, a target sample size of 350 participants was established. This comprised 150 participants with voice disorders, 150 vocally healthy control participants, and 50 participants for the test-retest reliability assessment. The final sample consisted of 150 participants with voice disorders, with 75 being Javanese speakers and 75 being Sundanese speakers. The Control Group also included 150 participants, age- and gender-matched to the Voice Disorder Group, with 75 Javanese speakers and 75 Sundanese speakers. The Test-Retest Reliability Group was composed of 50 participants, with 25 Javanese speakers and 25 Sundanese speakers.

The original English version of the VHI underwent a rigorous translation and cultural adaptation process to create Indonesian, Javanese, and Sundanese versions. This process adhered to established guidelines for cross-cultural adaptation of health-related instruments. The initial stage involved forward translation. Two independent translators, both of whom were fluent in English and the respective target language (Indonesian, Javanese, or Sundanese) and experienced in health-related translations, independently translated the VHI. One of the translators possessed a background in linguistics, while the other had a background in healthcare. Following the forward translation, a reconciliation process was undertaken. A third independent translator, who was also fluent in both English and the target language and had expertise in otorhinolaryngology, reviewed the two forward translations for each language. This translator's task was to create a reconciled version by resolving any discrepancies that existed between the two initial translations. The reconciled versions then underwent backward translation. Two different independent translators, who were native English speakers and were not aware of the original VHI, translated the reconciled versions back into English. An expert review panel was convened to evaluate the various translations. This panel consisted of otorhinolaryngologists, linguists, and cultural experts. The panel meticulously reviewed the original VHI, the forward translations, the reconciled versions, and the backward translations. The panel's objective was to assess the semantic, idiomatic, experiential, and conceptual equivalence of the translations. Any discrepancies or ambiguities identified during this review process were discussed in detail, and resolutions were reached through a process of consensus. Cognitive debriefing interviews were conducted to further refine the translated versions. These interviews involved a small sample of Javanese

and Sundanese speakers, with an equal number of individuals with and without voice disorders (n=10 per language). The purpose of these interviews was to evaluate the clarity, comprehensibility, and cultural relevance of the translated items. Participants were asked to paraphrase each item and to explain their understanding of the questions. Based on the feedback obtained from the cognitive interviews, minor adjustments were made to the wording of certain items to enhance their clarity and cultural appropriateness.

The data collection process involved administering the appropriate version of the I-VHI-JS (Indonesian, Javanese, or Sundanese) to each participant. This administration took place in a quiet room at the respective clinic. A trained research assistant was present during the data collection sessions to address any questions that participants had and to ensure that the questionnaire was completed properly. In addition to the I-VHI-JS, demographic and clinical data were collected from each participant. This included information on age, gender, education level, occupation, smoking status, and the duration of any voice problems. For participants in the Voice Disorder Group, the specific type of voice disorder was recorded based on the diagnosis provided by the otorhinolaryngologist. All participants in the Voice Disorder Group underwent a comprehensive otorhinolaryngological examination. This examination included indirect laryngoscopy or videostroboscopy, procedures used to confirm the diagnosis of the voice disorder and to assess its severity. Acoustic voice analysis was performed on all participants. A standardized voice recording protocol was employed to obtain voice samples. Participants were instructed to sustain the vowel /a/ for a minimum of three seconds, using a comfortable pitch and loudness. The recordings were conducted in a sound-treated booth, utilizing a high-quality microphone and a digital recording system. The acoustic analysis of the voice samples was carried out using specialized software (Praat). The acoustic parameters measured included; Jitter: This refers to the cycle-to-cycle variation in the fundamental frequency of the voice, expressed as a percentage; Shimmer: This refers to the cycle-to-cycle variation in the amplitude of the voice signal, also expressed as a percentage; Noise-to-Harmonics Ratio (NHR): This is the ratio of noise energy to harmonic energy within the voice signal; Fundamental Frequency (F0): This represents the basic frequency of the voice.

The statistical analysis of the data was performed using SPSS version 26 (IBM Corp., Armonk, NY). Descriptive statistics were used to summarize the demographic and clinical characteristics of the participants. These statistics included means, standard deviations, and frequencies. Internal consistency was assessed by calculating Cronbach's alpha coefficient for each subscale (Functional, Physical, and Emotional) and for the total score of each I-VHI-JS version. A Cronbach's alpha value of 0.70 or higher was considered to indicate acceptable internal consistency. Test-retest reliability was evaluated by calculating intraclass correlation coefficients (ICCs) with 95% confidence intervals for each I-VHI-JS version. An ICC value of 0.75 or higher was considered to indicate good reliability. Construct validity was examined by using independent samples t-tests to compare the I-VHI-JS scores (both total scores and subscale scores) between the Voice Disorder Group and the Control Group for each language version. Effect sizes, quantified as Cohen's d, were calculated to determine the magnitude of any observed differences between the groups. Concurrent validity was assessed by calculating Pearson correlation coefficients. These coefficients were used to examine the relationship between the I-VHI-JS scores (total and subscales) and the acoustic parameters (jitter, shimmer, NHR, and F0) for each language version. Discriminant validity was evaluated using receiver operating characteristic (ROC) curve analysis. This analysis was performed to assess the ability of the I-VHI-JS to accurately discriminate between individuals with voice disorders and those without voice disorders. The area under the curve (AUC) was calculated, with an AUC value of 0.80 or higher indicating excellent discrimination. The potential impact of demographic variables on I-VHI-JS scores was explored using two-way ANOVA. This analysis was conducted to investigate the interaction between language group (Javanese, Sundanese) and demographic variables

(age, gender, education level).

This study received approval from the Institutional Review Board of CMHC Indonesia. Prior to their participation in the study, written informed consent was obtained from all participants. Participants were given assurances regarding the confidentiality of their data and were informed of their right to withdraw from the study at any point.

3. Results

Table 1 presents a comprehensive overview of the demographic and clinical characteristics of the participants involved in the study. It provides a detailed comparison between the Voice Disorder Group, the Control Group, and the Test-Retest Group, with further stratification by Javanese and Sundanese speakers within each group; Age: The mean age of participants across all groups was around 42 years old. There were no significant differences in age between the Voice Disorder Group and the Control Group (p-value = 0.58), nor between the Javanese and Sundanese speakers within the Voice Disorder Group (p-value = 0.71). This indicates that the groups were well-matched in terms of age, reducing the likelihood of age being a confounding factor in the study's results; Gender: The gender distribution was approximately 60 males and 90 females in the Voice Disorder Group, and 62 males and 88 females in the Control Group. The Test-Retest Group had 20 males and 30 females. Statistical analysis showed no significant differences in gender distribution between the Voice Disorder and Control Groups (p-value = 0.72) or between Javanese and Sundanese speakers (p-value = 0.99). This suggests that gender was evenly distributed across the groups, minimizing its potential influence on the results; Language: As intended, each group was equally divided between Javanese and Sundanese speakers (75 in the Voice Disorder Group, 75 in the Control Group, and 25 in the Test-Retest Group). This balanced representation of the two language groups was crucial for the study's aim of developing and validating the I-VHI-JS for these specific populations; Education Level: The education levels were categorized into four groups: Primary, Junior Secondary, Senior Secondary, and Tertiary. The distribution of education levels was generally similar between the Voice Disorder and Control Groups, with the largest proportion of participants in the Tertiary education category. No significant differences were found in education level distribution between the Voice Disorder and Control Groups (p-value = 0.88), or between Javanese and Sundanese speakers. This suggests that the groups were comparable in terms of educational attainment, limiting the potential impact of education as a confounding variable; Occupation: Occupations were categorized into several groups, including Teacher, Singer, Call Center Operator, Office Worker, Healthcare Professional, Manual Laborer, and Other. There were significant differences in occupational distribution between the Voice Disorder and Control Groups (p-value < 0.001). The Voice Disorder Group had a higher proportion of teachers and singers compared to the Control Group. This is consistent with the understanding that these professions involve high vocal demands and are associated with a higher risk of voice disorders. The Control Group had a higher proportion of Office Workers. No significant differences were observed in occupational distribution between Javanese and Sundanese speakers (p-value = 0.95). The differences in occupational distribution between the Voice Disorder and Control Groups highlight the importance of considering occupational risk factors when studying voice disorders; Smoking Status: Smoking status was categorized as Never Smoker, Former Smoker, and Current Smoker. The Control Group had a significantly higher proportion of Never Smokers compared to the Voice Disorder Group (p-value = 0.002). Conversely, the Voice Disorder Group had a higher proportion of Former Smokers and Current Smokers. No significant differences were found in smoking status between Javanese and Sundanese speakers (p-value = 0.99). Among current smokers, the Voice Disorder Group reported a higher average number of cigarettes smoked per day (12.5) compared to the Control Group (8.2), with a statistically significant difference (p-value = 0.021). However, there was no significant difference between Javanese and Sundanese speakers (p-value=0.68). Smoking is a known risk factor for various health

problems, including respiratory and voice disorders. The differences in smoking status between the groups suggest that smoking may play a role in the development or exacerbation of voice disorders; Voice Disorder Type: This section provides a detailed breakdown of the specific voice disorder diagnoses within the Voice Disorder Group. The most common diagnoses were Muscle Tension Dysphonia (MTD) and Vocal Fold Nodules. This information is important for understanding the composition of the Voice Disorder Group and the range of voice pathologies included in the study; Duration of Voice Problems: The mean duration of voice problems in the Voice Disorder Group was 18.5 months. There was no significant difference in the duration of voice problems between Javanese and Sundanese speakers (p-value = 0.63). This data provides context regarding the chronicity of voice disorders experienced by the participants; Self-Reported Voice Severity: Voice severity was measured using a Visual Analog Scale (VAS) from 0 to 10. The Voice Disorder Group reported a significantly higher mean voice severity (7.2) compared to the Control Group (1.1) (p-value < 0.001). This confirms that the Voice Disorder Group had a substantially higher level of perceived voice impairment. No significant difference was found in self-reported voice severity between Javanese and Sundanese speakers (p-value = 0.58).

Table 1. Demographic and clinical characteristics of participants.

Characteristic	Voice Disorder Group (n=150)	Control Group (n=150)	Test-Retest Group (n=50)	p-value (VD vs. C)a	p-value (JS vs. SS)b
	Total (Jav / Sun)	Total (Jav / Sun)	Total (Jav / Sun)		
Age (years)	42.5 (12.8) (43.2 / 41.8)	41.8 (12.2) (42.5 / 41.1)	43.1 (13.0) (43.9 / 42.3)	0.58	0.71
Gender (Male/Female)	60/90 (30/45, 30/45)	62/88 (31/44, 31/44)	20/30 (10/15, 10/15)	0.72	0.99
Javanese Speakers (n)	75	75	25	-	-
Sundanese Speakers (n)	75	75	25	-	-
Education Level (n, %)					
- Primary (≤ 6 years)	20 (13.3%) (10, 10)	15 (10.0%) (7, 8)	7 (14.0%) (3, 4)	0.41	0.88
- Junior Secondary (7-9 years)	30 (20.0%) (15,15)	25 (16.7%) (12,13)	8 (16.0%) (4,4)		
- Senior Secondary (10-12 years)	25 (16.7%) (13, 12)	35 (23.3%) (18, 17)	10 (20.0%) (5, 5)		
- Tertiary (University/College)	75 (50.0%) (37, 38)	75 (50.0%) (38, 37)	25 (50.0%) (13, 12)		
Occupation (n, %)					
- Teacher	50 (33.3%) (25, 25)	10 (6.7%) (5, 5)	17 (34.0%) (8, 9)	<0.001	0.95
- Singer	20 (13.3%) (10, 10)	5 (3.3%) (2, 3)	7 (14.0%) (3, 4)	12	0.89
- Call Center Operator	10 (6.7%) (5, 5)	5 (3.3%) (3, 2)	3 (6.0%) (2, 1)	0.34	0.65
- Office Worker (Clerical/Admin)	25 (16.7%) (13, 12)	65 (43.3%) (32, 33)	9 (18.0%) (5, 4)	<0.001	0.91
- Healthcare Professional	5 (3.3%) (3, 2)	10 (6.7%) (5, 5)	2 (4.0%) (1, 1)	0.28	0.77
- Manual Laborer	15 (10.0%) (7, 8)	25 (16.7%) (13, 12)	6 (12.0%) (3, 3)	0.15	0.72
- Other (Student, Retired, Unemployed)	25 (16.7%) (12, 13)	30 (20.0%) (15, 15)	6 (12.0%)(3,3)	0.53	0.98
Smoking Status (n, %)					
- Never Smoker	90 (60.0%) (45, 45)	120 (80.0%) (60, 60)	30 (60.0%) (15, 15)	2	0.99
- Former Smoker	30 (20.0%) (15, 15)	15 (10.0%) (7, 8)	10 (20.0%) (5, 5)	48	0.99
- Current Smoker	30 (20.0%) (15, 15)	15 (10.0%) (8, 7)	10 (20.0%) (5, 5)	48	0.99
Cigarettes per day (Current Smokers)	12.5 (5.8) (12.1/12.9)	8.2 (4.3) (7.8 / 8.6)	13.1 (6.2) (12.5 / 13.7)	21	0.68
Voice Disorder Type (n, %)					
- Muscle Tension Dysphonia (MTD)	60 (40.0%) (30, 30)	-	-	-	-
- Primary MTD	35 (23.3%) (18, 17)	-	-		
- Secondary MTD	25 (16.7%) (12, 13)	-	-		
- Vocal Fold Nodules	45 (30.0%) (23, 22)	-	-	-	-
- Bilateral Nodules	30 (20.0%) (15, 15)	-	-	-	-
- Unilateral Nodules	15 (10.0%) (8, 7)	-	-	-	-
- Vocal Fold Polyps	23 (15.3%) (12, 11)	-	-	-	-
- Pedunculated Polyp	15 (10.0%) (8, 7)	-	-	-	-
- Sessile Polyp	8 (5.3%) (4, 4)	-	-	-	-
- Vocal Fold Paralysis	10 (6.7%) (5, 5)	-	-	-	-
- Unilateral Paralysis	8 (5.3%) (4, 4)	-	-	-	-
- Bilateral Paralysis	2 (1.3%) (1, 1)	-	-	-	-
- Vocal Fold Cysts	7 (4.7%) (4, 3)	-	-	-	-
- Other (Presbylaryngis, Laryngitis)	5 (3.3%) (2, 3)	-	-	-	-
Duration of Voice Problems (months)	18.5 (10.2) (19.1 / 17.9)	-	19.2 (9.8) (19.8 / 18.6)	-	0.63
Self-Reported Voice Severity (VAS 0-10)	7.2 (1.8) (7.4 / 7.0)	1.1 (0.9) (1.0 / 1.2)	7.3 (1.7) (7.5 / 7.1)	<0.001	0.58

[a]p-values are from independent samples t-tests for continuous variables and chi-square tests (or Fisher's exact test where appropriate) for categorical variables, comparing Voice Disorder vs. Control groups. [b]p-values comparing Javanese (Jav) and Sundanese (Sun) speakers within the Voice Disorder group using the same statistical tests. Data presented as mean (standard deviation) or n (%). Data within Javanese and Sundanese subgroups are presented in parentheses (Javanese / Sundanese). VAS = Visual Analog Scale.

Table 2 presents the results of the internal consistency analysis for the Indonesian Voice Handicap Index adapted for Javanese and Sundanese speakers (I-VHI-JS). Internal consistency, measured using Cronbach's alpha, assesses the extent to which the items within a scale or subscale are intercorrelated, indicating whether they consistently measure the same construct. The table demonstrates that all versions of the I-VHI-JS—Indonesian, Javanese, and Sundanese—exhibited high levels of internal consistency. The Cronbach's alpha values for the Total Score were 0.93 for the Indonesian version, 0.95 for the Javanese version, and 0.94 for the Sundanese version. These values are well above the commonly accepted threshold of 0.70, and even exceed 0.90, indicating excellent internal consistency for the overall I-VHI-JS across all three languages. This suggests that the items in each version of the I-VHI-JS are strongly intercorrelated and reliably measure the overall construct of voice handicap. The Cronbach's alpha values for the subscales—Functional (F), Physical (P), and Emotional (E)—also demonstrate strong internal consistency. For the Functional subscale, alpha values ranged from 0.88 to 0.91. For the Physical subscale, alpha values ranged from 0.90 to 0.92. For the Emotional subscale, alpha values ranged from 0.87 to 0.89. These results indicate that the items within each subscale are also highly intercorrelated, suggesting that each subscale consistently measures its respective domain of voice handicap (functional, physical, and emotional) across all three language versions.

Table 2. Internal consistency of the I-VHI-JS (Cronbach's Alpha).

I-VHI-JS Version	Total Score	Functional (F)	Physical (P)	Emotional (E)
Indonesian	0.93	0.88	0.90	0.87
Javanese	0.95	0.91	0.92	0.89
Sundanese	0.94	0.90	0.91	0.88

Table 3 presents the results of the test-retest reliability analysis for the Indonesian Voice Handicap Index adapted for Javanese and Sundanese speakers (I-VHI-JS). Test-retest reliability assesses the consistency of a measure over time, indicating whether it yields stable results when administered repeatedly to the same individuals under similar conditions. The Intraclass Correlation Coefficient (ICC) is the statistic used to measure this reliability. The table demonstrates that all versions of the I-VHI-JS—Indonesian, Javanese, and Sundanese—exhibited high levels of test-retest reliability. The ICC values for the Total Score were 0.88 for the Indonesian version, 0.90 for the Javanese version, and 0.89 for the Sundanese version. These values are well above the generally accepted threshold of 0.75 for good reliability, indicating excellent test-retest reliability for the overall I-VHI-JS across all three languages. This suggests that the I-VHI-JS provides stable and consistent measurements of voice handicap over time in Indonesian, Javanese, and Sundanese speakers. The ICC values for the subscales—Functional (F), Physical (P), and Emotional (E)—also demonstrate good to excellent test-retest reliability. For the Functional subscale, ICC values ranged from 0.82 to 0.86. For the Physical subscale, ICC values ranged from 0.85 to 0.88. For the Emotional subscale, ICC values ranged from 0.80 to 0.84. These results indicate that the items within each subscale also provide reasonably stable and consistent measurements of their respective domains of voice handicap (functional, physical, and emotional) over time across all three language versions.

Table 3. Test-retest reliability of the I-VHI-JS (Intraclass Correlation Coefficient - ICC).

I-VHI-JS Version	Total Score	Functional (F)	Physical (P)	Emotional (E)
Indonesian	0.88	0.82	0.85	0.80
Javanese	0.90	0.86	0.88	0.84
Sundanese	0.89	0.85	0.87	0.83

Table 4 presents the results of the construct validity assessment of the Indonesian Voice Handicap Index adapted for Javanese and Sundanese speakers (I-VHI-JS). Construct validity refers to the extent to which a test or measure accurately assesses the theoretical construct it is designed to measure. In this case, it examines whether the I-VHI-JS effectively differentiates between individuals with voice disorders and those without. The table compares I-VHI-JS scores between the Voice Disorder Group and the Control Group for the Indonesian, Javanese, and Sundanese versions of the instrument. For all three language versions (Indonesian, Javanese, and Sundanese), there were statistically significant differences in the Total Score between the Voice Disorder Group and the Control Group ($p < 0.001$). The Voice Disorder Group consistently exhibited much higher mean Total Scores than the Control Group. For the Indonesian version, the Voice Disorder Group had a mean Total Score of 68.5, while the Control Group had a mean of 22.1. For the Javanese version, the Voice Disorder Group had a mean Total Score of 72.3, and the Control Group had a mean of 24.5. For the Sundanese version, the Voice Disorder Group had a mean Total Score of 70.8, and the Control Group had a mean of 23.2. These results strongly indicate that the I-VHI-JS is able to distinguish between individuals with voice disorders and those without, supporting its construct validity. People with voice disorders perceive a greater handicap than people without voice disorders, as measured by the I-VHI-JS. Similar to the Total Scores, all three subscales (Functional, Physical, and Emotional) showed statistically significant differences ($p < 0.001$) between the Voice Disorder Group and the Control Group for all language versions. In each case, the Voice Disorder Group reported higher mean scores on all subscales compared to the Control Group. This indicates that individuals with voice disorders experience greater functional, physical, and emotional impacts related to their voice problems. The effect sizes, as measured by Cohen's d, were large across all language versions, ranging from 1.5 to 1.7. A Cohen's d of 0.8 or greater is generally considered a large effect. These large effect sizes demonstrate the substantial magnitude of the differences in I-VHI-JS scores between the Voice Disorder and Control Groups. This further supports the construct validity of the I-VHI-JS, showing that it effectively captures the significant differences in perceived voice handicap between individuals with and without voice disorders.

Table 4. Construct validity: comparison of I-VHI-JS scores between voice disorder and control groups.

I-VHI-JS Version	Group	Total Score (Mean ± SD)	Functional (F) (Mean ± SD)	Physical (P) (Mean ± SD)	Emotional (E) (Mean ± SD)	p-value	Cohen's d
Indonesian	Voice Disorder	68.5 ± 15.2	25.8 ± 6.5	22.3 ± 5.8	20.4 ± 5.1	<0.001	1.5
	Control	22.1 ± 8.7	8.2 ± 3.1	7.5 ± 2.9	6.4 ± 2.7		
Javanese	Voice Disorder	72.3 ± 14.8	27.1 ± 6.2	23.8 ± 5.5	21.4 ± 4.9	<0.001	1.7
	Control	24.5 ± 9.1	9.1 ± 3.3	8.2 ± 3.0	7.2 ± 2.8		
Sundanese	Voice Disorder	70.8 ± 15.5	26.5 ± 6.8	23.1 ± 6.0	21.2 ± 5.3	<0.001	1.6
	Control	23.2 ± 8.9	8.7 ± 3.2	7.8 ± 2.8	6.7 ± 2.9		

p-values are from independent samples t-tests.

Table 5 presents the results of the concurrent validity assessment of the Indonesian Voice Handicap Index adapted for Javanese and Sundanese speakers (I-VHI-JS). Concurrent validity examines the extent to which a measure correlates with other measures of the same construct that are administered at the same time. In this study, it assesses the relationship between the I-VHI-JS scores (a subjective measure of voice handicap) and acoustic parameters (objective measures of voice characteristics); Jitter Correlation: For all language versions (Indonesian, Javanese, and Sundanese), there were statistically significant positive correlations between Jitter (%) and the I-VHI-JS scores (Total Score and all subscales) ($p < 0.01$). The correlation coefficients ranged from 0.42 to 0.48 for the Total Score, 0.36 to 0.42 for the Functional subscale, 0.40 to 0.45 for the Physical subscale, and 0.33 to 0.39 for the Emotional subscale. These results indicate a moderate positive relationship between Jitter, a measure of vocal instability, and the perceived voice handicap as measured by the I-VHI-JS. Higher jitter values, indicating greater vocal instability, tend to be associated with higher VHI-JS scores, indicating a greater perceived voice handicap; Shimmer Correlation: Similar to Jitter, there were statistically significant positive correlations between Shimmer (%) and the I-VHI-JS scores (Total Score and all subscales) for all language versions ($p < 0.01$). The correlation coefficients ranged from 0.38 to 0.43 for the Total Score, 0.33 to 0.39 for the Functional subscale, 0.36 to 0.41 for the Physical subscale, and 0.30 to 0.36 for the Emotional subscale. These findings suggest a moderate positive relationship between Shimmer, another measure of vocal instability related to amplitude variation, and perceived voice handicap. Higher shimmer values are associated with higher VHI-JS scores; NHR Correlation: For all language versions, there were statistically significant negative correlations between Noise-to-Harmonics Ratio (NHR) and the I-VHI-JS scores (Total Score and all subscales) ($p < 0.01$). The correlation coefficients ranged from -0.35 to -0.28 for the Total Score, -0.30 to -0.23 for the Functional subscale, -0.32 to -0.26 for the Physical subscale, and -0.28 to -0.20 for the Emotional subscale. These results indicate a weak to moderate negative relationship between NHR, a measure of the proportion of noise in the voice signal, and perceived voice handicap. Higher NHR values, indicating more noise in the voice, tend to be associated with lower VHI-JS scores; F0 Correlation: The correlations between Fundamental Frequency (F0) and I-VHI-JS scores were generally weak and not statistically significant. For the Indonesian and Sundanese versions, none of the correlations between F0 and I-VHI-JS scores were statistically significant. For the Javanese version, there were statistically significant but weak negative correlations between F0 and Total Score (-0.18), Functional subscale (-0.14), Physical subscale (-0.16), and Emotional subscale (-0.11) ($p < 0.01$). These findings suggest that there is little to no consistent linear relationship between the fundamental frequency of the voice and the perceived voice handicap as measured by the I-VHI-JS.

Table 5. Concurrent validity: correlation between I-VHI-JS scores and acoustic parameters.

I-VHI-JS Version	Acoustic Parameter	Total Score	Functional (F)	Physical (P)	Emotional (E)
Indonesian	Jitter (%)	0.45	0.38	0.42	0.35
	Shimmer (%)	0.40	0.35	0.38	0.32
	NHR	-0.30	-0.25	-0.28	-0.22
	F0 (Hz)	-0.15	-0.10	-0.12	-0.08
Javanese	Jitter (%)	0.48	0.42	0.45	0.39
	Shimmer (%)	0.43	0.39	0.41	0.36
	NHR	-0.35	-0.30	-0.32	-0.28
	F0 (Hz)	-0.18	-0.14	-0.16	-0.11
Sundanese	Jitter (%)	0.42	0.36	0.40	0.33
	Shimmer (%)	0.38	0.33	0.36	0.30
	NHR	-0.28	-0.23	-0.26	-0.20
	F0 (Hz)	-0.12	-0.08	-0.10	-0.05

*$p < 0.01$, **$p < 0.05$.

Table 6 presents the results of the discriminant validity analysis of the Indonesian Voice Handicap Index adapted for Javanese and Sundanese speakers (I-VHI-JS). Discriminant validity assesses the ability of a measure to differentiate between groups that are theoretically expected to differ. In this case, it examines how well the I-VHI-JS can distinguish between individuals with and without voice disorders using Receiver Operating Characteristic (ROC) curve analysis; Area Under the Curve (AUC): The AUC values for the Total Score were 0.88 for the Indonesian version, 0.90 for the Javanese version, and 0.89 for the Sundanese version. The AUC values for the subscales ranged from 0.80 to 0.85 for the Indonesian version, 0.83 to 0.87 for the Javanese version, and 0.81 to 0.86 for the Sundanese version. AUC values range from 0.5 to 1.0, where 0.5 indicates no discrimination ability, and 1.0 indicates perfect discrimination. In this table, all AUC values are well above 0.80, indicating excellent discrimination ability of the I-VHI-JS and its subscales to distinguish between individuals with and without voice disorders across all three language versions; 95% Confidence Intervals (CI) for AUC: The 95% CIs for all AUC values are relatively narrow and do not include 0.5, further supporting the statistical significance and reliability of the discrimination ability; p-value (AUC): All p-values for the AUC were < 0.001, indicating that the discrimination ability of the I-VHI-JS is highly statistically significant; Optimal Cut-off Score: The table provides optimal cut-off scores for the Total Score and each subscale for each language version. These cut-off scores can be used in clinical practice to screen individuals for voice disorders. Scores above the cut-off suggest a higher likelihood of having a voice disorder. For example, the optimal cut-off score for the Total Score is 35 for the Indonesian version, 38 for the Javanese version, and 36 for the Sundanese version; Sensitivity: Sensitivity refers to the ability of the test to correctly identify individuals with a voice disorder. Sensitivity values ranged from 88% to 90% for the Total Score and 80% to 87% for the subscales. These high sensitivity values indicate that the I-VHI-JS is effective at identifying most individuals who truly have a voice disorder; Specificity: Specificity refers to the ability of the test to correctly identify individuals without a voice disorder. Specificity values ranged from 85% to 87% for the Total Score and 78% to 84% for the subscales. These high specificity values indicate that the I-VHI-JS is effective at correctly identifying most individuals who do not have a voice disorder; Positive Predictive Value (PPV): PPV is the probability that an individual *with* a positive test result (above the cut-off) truly has a voice disorder. PPV values ranged from 86% to 88% for the Total Score and 79% to 85% for the subscales; Negative Predictive Value (NPV): NPV is the probability that an individual with a negative test result (below the cut-off) truly does not have a voice disorder. NPV values ranged from 87% to 89% for the Total Score and 79% to 86% for the subscales; Youden's Index (J): Youden's Index is a measure of the overall diagnostic effectiveness of a test, combining both sensitivity and specificity. Values range from 0 to 1, with higher values indicating better performance. Youden's Index values ranged from 0.73 to 0.77 for the Total Score and 0.58 to 0.71 for the subscales, indicating good overall diagnostic effectiveness; Accuracy: Accuracy is the overall proportion of individuals who are correctly classified by the test. Accuracy values ranged from 86.5% to 88.5% for the Total Score and 79.0% to 85.5% for the subscales, indicating high overall accuracy of the I-VHI-JS.

Table 7 (it's important to note that this is the second Table 7 in the set of images) presents the results of a two-way ANOVA examining the influence of demographic variables and their interaction with language group on the I-VHI-JS scores. ANOVA, or Analysis of Variance, is a statistical test used to compare means between groups. In this case, it investigates whether age, gender, and education level, and their interactions with language group (Javanese vs. Sundanese speakers), have a statistically significant effect on the I-VHI-JS scores. The F-value for age is 1.21, and the p-value is 0.28. Since the p-value is greater than the significance level of 0.05, there is no statistically significant main effect of age on the I-VHI-JS scores. This suggests that age, considered alone, does not significantly influence the

I-VHI-JS scores. The F-value for gender is 0.85, and the p-value is 0.36. The p-value is greater than 0.05, indicating that there is no statistically significant main effect of gender on the I-VHI-JS scores. This means that gender, by itself, does not significantly affect the I-VHI-JS scores. The F-value for education level is 2.12, and the p-value is 0.13. With a p-value greater than 0.05, there is no statistically significant main effect of education level on the I-VHI-JS scores. This suggests that differences in education level, when considered alone, do not significantly impact the I-VHI-JS scores. The F-value for language group is 0.54, and the p-value is 0.47. The p-value is greater than 0.05, indicating that there is no statistically significant main effect of language group on the I-VHI-JS scores. This means that being a Javanese speaker or a Sundanese speaker, considered alone, does not significantly influence the I-VHI-JS scores. The F-value for the interaction between age and language group is 0.92, and the p-value is 0.40. The p-value is greater than 0.05, showing that there is no statistically significant interaction effect between age and language group on the I-VHI-JS scores. This suggests that the effect of age on I-VHI-JS scores does not differ significantly between Javanese and Sundanese speakers. The F-value for the interaction between gender and language group is 1.35, and the p-value is 0.25. The p-value is greater than 0.05, indicating that there is no statistically significant interaction effect between gender and language group on the I-VHI-JS scores. This implies that the effect of gender on I-VHI-JS scores does not significantly differ between Javanese and Sundanese speakers. The F-value for the interaction between education level and language group is 0.77, and the p-value is 0.55. The p-value is greater than 0.05, demonstrating that there is no statistically significant interaction effect between education level and language group on the I-VHI-JS scores. This suggests that the effect of education level on I-VHI-JS scores does not significantly differ between Javanese and Sundanese speakers.

Table 6. Discriminant validity of the I-VHI-JS: ROC curve analysis and derived metrics.

I-VHI-JS Version	Metric	Total Score	Functional (F)	Physical (P)	Emotional (E)
Indonesian	Area Under the Curve (AUC)	0.88	0.85	0.82	0.80
	95% CI for AUC	0.84-0.92	0.81-0.89	0.78-0.86	0.76-0.84
	p-value (AUC)	<0.001	<0.001	<0.001	<0.001
	Optimal Cut-off Score	35	15	13	11
	Sensitivity	88%	85%	82%	80%
	Specificity	85%	82%	80%	78%
	Positive Predictive Value (PPV)	86%	83%	81%	79%
	Negative Predictive Value (NPV)	87%	84%	81%	79%
	Youden's Index (J)	0.73	0.67	0.62	0.58
	Accuracy	86.5%	83.5%	81.0%	79.0%
Javanese	Area Under the Curve (AUC)	0.90	0.87	0.85	0.83
	95% CI for AUC	0.87-0.93	0.83-0.91	0.81-0.89	0.79-0.87
	p-value (AUC)	<0.001	<0.001	<0.001	<0.001
	Optimal Cut-off Score	38	17	15	13
	Sensitivity	90%	87%	85%	83%
	Specificity	87%	84%	82%	80%
	Positive Predictive Value (PPV)	88%	85%	83%	81%
	Negative Predictive Value (NPV)	89%	86%	84%	82%
	Youden's Index (J)	0.77	0.71	0.67	0.63
	Accuracy	88.5%	85.5%	83.5%	81.5%
Sundanese	Area Under the Curve (AUC)	0.89	0.86	0.84	0.81
	95% CI for AUC	0.85-0.93	0.82-0.90	0.80-0.88	0.77-0.85
	p-value (AUC)	<0.001	<0.001	<0.001	<0.001
	Optimal Cut-off Score	36	16	14	12
	Sensitivity	89%	86%	84%	81%
	Specificity	86%	83%	81%	79%
	Positive Predictive Value (PPV)	87%	84%	82%	80%
	Negative Predictive Value (NPV)	88%	85%	83%	80%
	Youden's Index (J)	0.75	0.69	0.65	0.60
	Accuracy	87.5%	84.5%	82.5%	80.0%

Table 7. Two-way ANOVA result of demographic variable.

Variable	F Value	p-value
Age	1.21	0.28
Gender	0.85	0.36
Education Level	2.12	0.13
Language Group	0.54	0.47
Age * Language Group	0.92	0.40
Gender * Language Group	1.35	0.25
Education Level * Language Group	0.77	0.55

4. Discussion

The rigorous translation and cultural adaptation process, following established guidelines, ensured that the I-VHI-JS is linguistically and culturally appropriate for Indonesian, Javanese, and Sundanese speakers. The forward-backward translation, expert review, and cognitive debriefing steps were crucial for identifying and resolving potential sources of bias and ensuring that the translated items accurately reflected the meaning and intent of the original VHI. This meticulous process is essential because direct translations of PROMs without considering cultural nuances can compromise the validity and reliability of the instrument. Indonesia, with its diverse linguistic landscape, necessitates careful adaptation to capture the subtleties of language and cultural expressions related to voice handicap. The success of the adaptation process in this study supports the notion that cultural adaptation in the development of PROMs is of paramount importance. Failing to account for cultural differences can lead to inaccurate assessments and misinterpretations of patient experiences.[11-14]

The high internal consistency (Cronbach's alpha > 0.90) and test-retest reliability (ICC > 0.85) of all I-VHI-JS versions indicate that the instrument is stable and consistent in measuring voice handicap. The Cronbach's alpha values for the Total Score were 0.93 for the Indonesian version, 0.95 for the Javanese version, and 0.94 for the Sundanese version. These values are well above the commonly accepted threshold of 0.70, and even exceed 0.90, demonstrating excellent internal consistency for the overall I-VHI-JS across all three languages. This suggests that the items in each version of the I-VHI-JS are strongly intercorrelated and reliably measure the overall construct of voice handicap. The Cronbach's alpha values for the subscales—Functional (F), Physical (P), and Emotional (E)—also demonstrate strong internal consistency. For the Functional subscale, alpha values ranged from 0.88 to 0.91. For the Physical subscale, alpha values ranged from 0.90 to 0.92. For the Emotional subscale, alpha values ranged from 0.87 to 0.89. These results indicate that the items within each subscale are also highly intercorrelated, suggesting that each subscale consistently measures its respective domain of voice handicap (functional, physical, and emotional) across all three language versions. The Intraclass Correlation Coefficient (ICC) values for the Total Score were 0.88 for the Indonesian version, 0.90 for the Javanese version, and 0.89 for the Sundanese version. These values are well above the generally accepted threshold of 0.75 for good reliability, indicating excellent test-retest reliability for the overall I-VHI-JS across all three languages. This suggests that the I-VHI-JS provides stable and consistent measurements of voice handicap over time in Indonesian, Javanese, and Sundanese speakers. The ICC values for the subscales—Functional (F), Physical (P), and Emotional (E)—also demonstrate good to excellent test-retest reliability. For the Functional subscale, ICC values ranged from 0.82 to 0.86. For the Physical subscale, ICC values ranged from 0.85 to 0.88. For the Emotional subscale, ICC values ranged from 0.80 to 0.84. These results indicate that the items within each subscale also provide reasonably stable and consistent measurements of their respective domains of voice handicap (functional, physical, and emotional) over time across all three language versions. These findings are comparable to those reported in previous VHI validation studies in other languages, supporting the

cross-cultural robustness of the VHI construct. The strong internal consistency and test-retest reliability across all versions of the I-VHI-JS validate its stability and consistency in measuring voice handicap. This means that the items within the I-VHI-JS are measuring the same underlying construct and that the instrument is reliable in its measurement of voice handicap in Indonesian, Javanese, and Sundanese speakers. The findings support the use of the I-VHI-JS as a reliable tool for assessing voice handicap in these populations and for monitoring changes in voice handicap over time.[15-20]

5. Conclusion

In conclusion, this study successfully developed and validated the Indonesian Voice Handicap Index adapted for Javanese and Sundanese speakers (I-VHI-JS). The rigorous translation and cultural adaptation process ensured the instrument's linguistic and cultural appropriateness. The I-VHI-JS demonstrated excellent internal consistency and test-retest reliability, indicating its stability and consistency in measuring voice handicap. The instrument also showed strong construct validity, effectively differentiating between individuals with and without voice disorders. Concurrent validity was supported by moderate correlations between I-VHI-JS scores and acoustic parameters. Furthermore, the I-VHI-JS demonstrated excellent discriminant validity, accurately distinguishing between individuals with and without voice disorders. The findings indicate that the I-VHI-JS is a reliable and valid tool for assessing voice-related handicaps in Indonesian, Javanese, and Sundanese speakers. It can be used in clinical practice and research to evaluate the impact of voice disorders and monitor treatment outcomes. The availability of a culturally appropriate instrument is crucial for accurate assessment and effective management of voice disorders in these populations.

6. References

1. Batista D de J, da Silva RD, Martins AA, de Araújo CM, Santos RS, Filho OG, et al. Internal consistency of the Voice Handicap Index in individuals with dysphonia: a systematic review and meta-analysis. J Voice. 2023.
2. Huang M, Yu L, Hu J, Ren J, Li Z, Pan Z, et al. Adaptation and validation of the Mandarin Chinese version of the pediatric Voice Handicap Index-10 (pVHI-10). J Voice. 2024; 38(1): 243.e31-243.e35.
3. Cialente F, Torsello M, Meucci D, Tropiano ML, Salvati A, Trozzi M. Pediatric voice handicap index (pVHI): A tool for evaluating the reliability and validity of voice therapy in children with benign vocal fold nodules. J Voice. 2024; 38(6): 1527.e21-1527.e26.
4. Adriaansen A, Van Lierde K, Meerschman I, Everaert C, D'haeseleer E. Validity and reliability of the Dutch children's voice handicap index-10. J Voice. 2024; 38(6): 1528.e11-1528.e19.
5. Yu L, Ren J, Xu T, Ma T, Pan Z, Lu D. A comparative and correlational study of the vocal fatigue index (VFI) and the voice fatigue handicap questionnaire (VFHQ) in individuals with and without voice disorders. J Voice. 2024; 38(6): 1529.e1-1529.e6.
6. Baghban K, Seifpanahi M-S, Nazari M-A, Hashemnia SS, Manoosi M, Moshtagh A. Cross-cultural adaptation and validation of the Persian version of the children's voice handicap index-10 for parents. J Voice. 2024.
7. El-Deeb ME, Elzayat S, El-Sobki A, Salamah A, Gehad I, Piazza C. Voice handicap index-throat: Translation and cross-cultural adaptation to the Arabic language. J Voice. 2024.
8. Okui A, Nimura Y, Komazawa D, Kanazawa T, Konomi U, Hirosaki M, et al. Validation of the Japanese version of the singing voice handicap index. J Voice. 2024; 38(2): 544.e15-544.e21.
9. Sari A, Esen Aydinli F, Incebay O, Ozcebe E, Cadallı Tatar E, Barmak E, et al. Development of a voice handicap index for clergymen: an investigation of its validation and reliability. J Voice. 2024.

10. Harati O, Martineau S, Paquin A, Brisebois S, Martel-Sauvageau V, Leclerc A-A. Quebec French translation, cultural adaptation, and validation of the singing voice handicap index-10 questionnaire for singers with dysphonia. J Voice. 2024
11. Hosseiniheshmatian M, Darouie A, Hasanvand A, Bakhshi E. A comparison of the dysphonia severity index and singing voice handicap index between trained and untrained male Iranian traditional singers. J Voice. 2024.
12. Ribeiro VV, Batista D de J, de Castilho WLS, da Silva IKLB, Casmerides MCB, do Carmo RD, et al. Reliability, measurement error, and responsiveness of the voice handicap index: a systematic review and meta-analysis. J Voice. 2024.
13. Nazari M-A, Saeedi S, Seifpanahi M-S, Hadi N. Adaptation and validation of the Persian version of the voice handicap index-10 (VHI-10). J Voice. 2024.
14. De Maio V, Mozzanica F, Cardella A, Schindler A, Murry T, Ricci-Maccarini A. Normative values for the children voice handicap index-10 (CVHI-10) and for the children voice handicap index-10 for parents (CVHI-10-P). J Voice. 2024.
15. Doğan N, Eser MT, Yağcıoğlu D, Esen Aydınlı F. A meta-analytic reliability generalization study of the pediatric voice handicap index. J Voice. 2024.
16. Šimkienė R, Šiupšinskienė N, Purlys P, Žvirblis T, Lesinskas E. Adaptation and validation of the Lithuanian version of the pediatric voice handicap index. J Voice. 2024; 38(5): 1247.e25-1247.e30.
17. Fu S, Chen Y-C, Wang C-T. Reliability and validity of the Mandarin version of singing voice handicap index - 10 (MSVHI-10). J Voice. 2025; 39(1): 279.e1-279.e5.
18. Giannopoulou J, Papadopoulou E, Bibas A, Papathanasiou I. Validation of the singing voice handicap index in Greek singers: Normal and voice-disordered participants. J Voice. 2025.
19. Azadnajafabad S, Mohammadi E, Afzal S, Sohrabi H, Dabirmoghaddam P, Aghazadeh K. Clinical characteristics and Voice Handicap Index assessment of common benign vocal fold lesions: a case-control study from a referral center. J Voice. 2025.
20. Ello A, Sõber L, Viigand M, Lehes L. Cross-cultural adaptation and validation of the voice handicap index into Estonian. J Voice. 2025.